IMPLICIT AND DECLARATIVE LEARNING AND MATHEMATICS INSTRUCTION

by

Timothy C. Indahl

A dissertation submitted to the faculty of
The University of Utah
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

Department of Educational Psychology

The University of Utah

May 2015

The University of Utah Graduate School

STATEMENT OF DISSERTATION APPROVAL

The dissertation of **Timothy C. Indahl**

has been approved by the following supervisory committee members:

Dan Woltz	, Chair	**11/14/2014** Date Approved
Michael Gardner	, Member	**11/14/2014** Date Approved
Anne Cook	, Member	**11/14/2014** Date Approved
Mary Burbank	, Member	**11/14/2014** Date Approved
Kirsten Butcher	, Member	**11/14/2014** Date Approved

and by **Anne Cook** , Chair/Dean of

the Department/College/School of **Educational Psychology**

and by David B. Kieda, Dean of The Graduate School.

ABSTRACT

A review of traditional mathematics instruction suggests that conventional methods emphasize student learning of declarative rules about math problem solving procedures, especially in the early phase of learning. In contrast to the implicit learning of procedural skills, this approach places heavy demands on working memory and may be partly responsible for low levels of math achievement by many students. The present study explored the plausibility of implicit learning of polynomial problem structure prior to declarative rule instruction and its impact on subsequent problem solving skill, rule learning, and perception of difficulty. Participants selected proper factorizations of quadratic polynomials from two possible answer choices over many blocks in a task that was structured to achieve errorless learning through a vanishing cues approach. Measures were administered to assess problem solving skill, rule understanding, and perception of learning difficulty. Evidence supports the hypothesis that some mathematics skill can be learned implicitly, but marginal and conflicting results raise questions about the impact of initial implicit learning on subsequent rule learning and difficulty perception. Findings are interpreted with respect to implicit learning and skill acquisition theories.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Figure

CHAPTER 1

INTRODUCTION

Mathematics is a school subject that is highly scrutinized in our society today. Students' ability to understand and apply mathematical concepts in day-to-day life and work is considered essential for success in adulthood. Despite the critical priority often assigned to mathematics teaching, learning, and research, many students still struggle with basic concepts and problem solving.

There have been a multitude of studies, using both experimental and action research methods, which have attempted to identify the best way to teach mathematics to children. Many of these have been successful in that they identify methods that produce a positive benefit to student learning, but even the most promising instructional methods yield some students who struggle to learn. Furthermore, the traditionally used textbooks, lecture, and homework methods found in most public schools tend to result in many students who struggle to learn, and later apply, complex mathematical concepts and problems solving skills.

The National Center for Education Statistics reports the findings of the National Assessment of Educational Progress, a periodically administered assessment used to evaluate student proficiency throughout the United States in mathematics and other content areas. According to the National Center for Education (2011) only 35% of students in the nation were assessed as proficient in 8th grade mathematics (National Center for Education, 2011).

Given that a national assessment suggests approximately two-thirds of students entering high school in the United States lack proficiency in mathematics as defined by the National Center for Education, it is vital to investigate the root of student struggles in this area. In mathematics, it is traditional protocol to begin by teaching students the rules for a concept or problem solution and then to have students apply the rules in practice. This methodology of declarative rule learning preceding practice is common in mathematics education in the United States, and given the limitations of human working memory, could possibly contribute to the difficulties many students face when learning math. In order to acquire math problem-solving skills in this way, a student needs to learn the rules for solving a problem and then hold those rules in working memory while trying to apply the rules. Working memory is limited in its capacity for holding new, unfamiliar information. Given the complex rule structure of many math skills, and the use of novel symbols to represent variables and operations, the initial declarative understanding of verbal descriptions for new math skills can place an enormous strain on students' working memory. An important goal, then, is to discover if there are methods to reduce the initial demands on learners' working memory that may be partially responsible for the frequency of poor math achievement in our schools.

One approach to minimizing the impact of working memory limits on students' acquisition of new math skills is to initially promote an implicit rather than explicit understanding of the operations to be learned. A common example of implicit acquisition of complex rules is the way young children acquire the ability to comprehend and speak their native language. Infants are equipped with extremely limited working memory abilities and no declarative knowledge about the language they will learn. However, by the time an average child is 5 or 6 years old, they are able to communicate fluently in their native language. What is more, the grammar rules that govern most languages are very complex, yet children are typically able to communicate with very few grammatical errors. Children are not taught, explicitly, the rules that govern their language or the proper way to apply those rules. Instead, children are nearly continuously exposed to people speaking their language. Over time, this exposure results in children learning the rules underlying that language, that is, the grammar of the language. They are able to initially understand what other people around them are saying, and eventually are able to speak themselves, all the time applying the grammar rules they are learning implicitly through exposure.

The way children learn the grammar and other elements of a language is an example of implicit learning of procedural knowledge. It is implicit because it occurs largely in the absence of conscious, effortful processes and results in a procedural memory that can be used in practice, even though that information is not explicitly available as declarative knowledge. This procedural memory is not only useful in children's language production; it makes it easier for them to learn the grammar rules declaratively when they are older. They have an implicit sense as to whether statements and sentences are grammatically correct, which is invaluable when trying to explicitly understand complex grammatical rule definitions.

Can this example of implicit grammar learning provide a key to alleviating the constraint working memory imposes on learning mathematics problem solving skills? Although it is possible that language acquisition is a special case in implicit learning, there is evidence suggesting that young children possess well-developed implicit memory functions in domains other than grammar (Parkin, 1997). The research reported here investigates the effectiveness of initial exposure to implicit learning opportunities prior to declarative instruction in the domain of mathematics. If students initially implicitly learn a "grammar" of mathematics problems within a domain, they might develop a form of procedural knowledge for solving that type of problem. If this is possible, the implicit knowledge might exist and function without the declarative knowledge that places high demands on working memory in early stages of learning. Just as implicit knowledge of grammar helps a child later learn the declarative knowledge of grammatical rules, the implicit procedural knowledge of math problems could potentially facilitate the subsequent explicit understanding of procedures for solving those problems.

The present study compares algebraic problem solving skill learning between two groups that completed Implicit and Declarative Learning Tasks, but in different orders. Participants completed training in their assigned condition, implicit-first or declarative-first, during Sessions 1 and 2, and then completed the other type of training during Session 3. In addition to the learning tasks, participants completed several tasks over three sessions designed to measure skill acquisition and transfer, declarative rule learning, and participant perception of difficulty for both implicit and Declarative Learning Tasks.

CHAPTER 2

LITERATURE REVIEW

Procedural and declarative memory are viewed by many theorists as separate systems for learning and retaining skills and knowledge (Anderson, 1993; Cohen & Squire, 1980; Nissen, Knopman, & Schacter, 1987). Each of these types of memory has been studied extensively over the past several decades. In this chapter I review some of the history of skill acquisition research, with a focus on implicit procedural learning as well as research about interactions between procedural and declarative learning, artificial grammar, working memory, and errorless learning. In addition, I review the literature for best and current practices in mathematics instruction, including common mathematics textbooks.

Learning and Skill Acquisition

John R. Anderson (1982) introduced the adaptive control of thought (ACT) theory of cognition, which described learning processes and served as a model for how skill acquisition occurs. His model described two components to learning, both a conscious declarative part and an 'unconscious' procedural part. Anderson's model posited that skill learning is a two-stage process and that declarative learning must be the first stage. The model required declarative processes for the interpretation of facts, which were necessary before the mind could form procedures (if-then production rules) for whatever skill was being learned. Anderson stated, "In the first stage the learner receives instruction and information about a skill. The instruction is encoded as a set of facts about the skill. These facts can be used by general interpretive procedures to generate behavior" (Anderson, 1982, p. 370). Thus, the relevant information from Anderson's model is that procedural skill learning must be preceded by an initial declarative processing phase.

Anderson updated his model of skill acquisition in the theory of cognition known as ACT-R (1982, 1993). His model maintained the idea that "all knowledge starts out in declarative form" (Anderson, 1993, p. 69). The ACT-R model posits that learning of procedural skills is mediated by declarative knowledge using an analogy mechanism. This mechanism helps form production rules, from declarative knowledge, which are essentially mental procedures that will activate given the proper circumstances. In response to evidence of procedural learning in absence of declarative learning by amnesiacs

(Willingham, Nissen, & Bullemer, 1989), Anderson suggested that declarative learning does occur first in these impaired individuals, but that it isn't encoded strongly enough to be remembered. To that end, Anderson stated, "the best interpretation of amnesia in ACT-R is in terms of weak initial memory traces" (Anderson, 1993, p. 25).

In a further update to ACT-R, Anderson and colleagues (2004), did not specifically state that all procedural knowledge must first pass through a declarative form, but their learning experiments all required participants to read and memorize instructions for a task, with the intention that those instructions would become proceduralized with significant practice. His computational model also indicates that declarative instruction knowledge must be retrieved from memory early in the skill acquisition process (Anderson et al., 2004). Thus, Anderson's updated model of skill acquisition continues to promote the importance of declarative knowledge in the formation of procedural skills.

Implicit Learning in Absence of Explicit Learning

While Anderson's model indicates that learning procedural skills requires initial declarative processing, research by others in the field suggests that learning could occur in the absence of declarative memory processes. Willingham, Nissen and Bullemer (1989) published a study utilizing the serial response task, or SRT, in order to study this effect. This task measures time to respond to one of four keys after a display signal indicates which key to press. In the critical task condition, each block of continuous trials contains a repeating sequence of 12 key presses. Participants were able to perform with shorter response times (RT) after practicing when a repeating sequence was present compared to a random sequence condition (Willingham et al., 1989). Importantly, participants were not told a sequence existed but were able to implicitly learn the pattern without declaratively learning the sequence first. In fact, there are studies that show implicit learning in this task provides for better performance than explicitly learning the same sequence (Reber & Squire, 1998).

Further evidence of the ability to implicitly learn procedural skills can be seen through learning studies with amnesic patients and patients with other memory disorders that degrade declarative learning ability. Such patients were able to learn procedural sequences despite being unable to learn declaratively (Willingham et al., 1989). A separate study of amnesic and Korsakoff syndrome participants showed similar learning of procedural skills in the absence of declarative learning (Fahle & Daum, 2002). In both studies, participants provided verbal reports that indicated they had no declarative memories of their respective tasks. In addition, "amnesiac patients show improvement in problem solving and learning to operate complex equipment despite no conscious memory of the training" (Litman & Reber, 2005, p. 441). These cases support the idea that individuals can learn procedural skills implicitly, with little or no access to declarative knowledge of those skills.

Over the course of several studies, Nissen further explored the idea of implicit learning in order to better understand the role of attention and conscious awareness in procedural skill learning. In one study, participants performed the SRT under different conditions. Those participants who were exposed to a divided attention task did not learn the sequence, as evidenced by the lack of facilitation in SRT responses when the repeating sequence was present (Nissen, 1992; Nissen & Bullemer, 1987). Nissen concluded "learning the sequence required attentional capacity but not awareness" (Nissen, 1992, p. 206). This research suggests a difference between explicit awareness of task information and attention to the task, and this conclusion has been supported by others (Corr, 2003; Hartman, Knopman, & Nissen, 1989). The conclusion suggests that learners must be attending to a procedural learning task in order to implicitly acquire the intended skills, but that this attention is separate from declarative knowledge of the skill rules because the latter is not required for skill learning.

Interactions between implicit and explicit knowledge

Some evidence suggests there are facilitative interactions between declarative and procedural memory processes. For example, Willingham, Nissen, and Bullemer's (1989) SRT study found that a subset of participants could use the procedural knowledge they had learned to generate a declarative understanding of the sequences, in complete absence of declarative training. Those participants who spontaneously acquired a declarative knowledge of the repeating sequence also showed greater RT facilitation than those who did not. In order to understand this more fully, the researchers attempted to control for anticipatory responses (responding prior to stimulus presentation) from participants with declarative pattern knowledge. They reasoned that correct responses of less than 100 ms did not allow for processing the stimuli, making a choice and responding, so they were indicative of participants anticipating the next stimulus position. The authors posited that those who induced a declarative understanding of the pattern used that knowledge to anticipate the next stimulus position. When they removed these anticipatory responses from the analysis, the facilitation advantage shown by those who acquired declarative knowledge was eliminated; these participants now showed equivalent facilitation patterns as those who did not induce a declarative understanding. It is not entirely clear when or how declarative learning occurred, but their findings suggest that induction of declarative rules is possible under implicit learning conditions that lack explicit instruction.

Further supporting a positive relationship between implicit and explicit learning, an electroencephalography (EEG) study showed spontaneous declarative learning of procedurally learned patterns (Wessel, Haider, & Rose, 2012). In this study, participants performed an SRT in which a pattern was present in the responses. Some of the participants became declaratively aware of the pattern during the task and the EEG of those patients showed a corresponding change that was not present in the participants who reported being unaware of the pattern. According to the authors, "we found changes in high-frequency gamma-band EEG coherence in the rPFC to be associated with the transition between implicit and explicit contingency awareness in explicit learners in a serial reaction time task" (Wessel et al., 2012, p. 161). This finding provided biological evidence for the possibility of procedurally learned skills preceding, and providing the basis for, declarative knowledge of the skill. This and the work of Willingham are contrary to some current theory on the matter. Nevertheless, if substantiated with additional research, such findings have potentially important implications for initial implicit learning of complex skills with declarative rules that can overload working memory and impede early phases of skill acquisition.

Grammar and Artificial Grammar

Implicit learning has been widely studied through the use of the artificial grammar paradigm. Artificial grammar learning involves participants being directed to remember strings of letters that were generated by artificial, grammar-like rules. An early study by Reber (1967) showed that simply studying these letter strings without exposure to the grammar rules could lead to recognition of novel strings created with the same rules.

In a later study of artificial grammar, Dienes, Broadbent, and Berry (1991) also demonstrated that participants can learn artificial rules of grammar for strings of letters without any direct declarative processing of the rules. Participants were simply exposed to strings of letters and instructed to memorize them. In a later task, they were able to accurately identify nearly two-thirds of new letter strings that followed the grammar rules of the strings they had memorized. Given that participants could not effectively explain how they knew test items followed the rules in a free report task, this experiment provided further evidence that implicit learning of an underlying rule structure does not require direct declarative processing of those rules. Furthermore, Dienes and colleagues told some participants that there were grammar rules present in the letter strings and the participants should attempt to learn them; these participants showed no advantage in identifying new letter strings that followed the grammar rules compared to the group that was only instructed to memorize the strings (Dienes et al., 1991). In this case, not only was declarative processing of the underlying rules not required for procedural learning, instructions to focus on explicit rule learning did not provide any benefit.

Artificial grammar research has provided experimental evidence that people can learn to use a complex set of rules without learning them declaratively. This may be similar to what occurs when children learn the grammar of their native language in speaking, and later writing. They acquire the ability to apply complex rules of grammar without an explicit, declarative understanding of those rules. It is important to consider, however, the possibility that language acquisition may be unique in the implicit learning domain, as people may simply be innately equipped to learn a language and its grammar. If this is the case, then implicit learning in other domains may be more difficult to promote; if not, the artificial grammar and other implicit learning research supports the idea that some forms of pattern learning can be implicit in nature, and that in some cases, declarative understanding might have little benefit to performance.

Working Memory and Individual Differences

Current theory suggests that most forms of declarative learning depend on working memory resources and many studies have documented decreased learning, problem solving, and knowledge retrieval when working memory resources are taxed (Anderson, Reder, & Lebiere, 1996; Ashcraft & Krause, 2007). In contrast, there is evidence suggesting that implicit learning may not put the same demands on working memory. Research in skill acquisition has shown that factors such as working memory (Woltz, 1988) and general intelligence (Ackerman, 1988, 1992) have greater relationships to early skill learning performance when declarative processes are required and weaker relationships when skills have approached a level of procedural automaticity.

Measures of general intelligence (e.g., IQ scores) are often used to predict the ability of people to learn novel skills, but do they also predict implicit learning ability? In a study comparing implicit and explicit learning, Reber, Walkenfeld, and Hernstadt (1991) investigated that question. Participants in the study performed both an implicit and an explicit learning task and were given a measure of intelligence. The results showed that while participants' performance indicated many individual differences on the explicit learning task, there were very few individual differences on the Implicit Learning Task. Likewise, IQ scores correlated highly with the explicit learning task, but not with the implicit task (Reber, 1991). The authors interpreted these findings broadly to suggest "evolutionarily and phylogenetically older implicit processes ought to show a tighter distribution of performance than the more recently emergent explicit processes" (Reber, 1991, p. 894).

The low variability in performance of implicit memory across people predicted by A. S. Reber and colleagues (1991) would likely also support smaller effects of working memory limits on the use and formation of memory utilizing implicit processes. A study by P. J. Reber and Kotovsky (1997) studied the effect of working memory demands on implicit learning through a problem-solving task, but their findings contradict this prediction. Participants were presented the balls and boxes task, a puzzle in which five balls were located in boxes and had to be removed. A set of rules governed when a ball could be moved in and out of its box. Their results indicated that additional working memory demands had a negative effect on implicit problem solving, as evidenced by increased difficulty when initially solving the balls and boxes task while under cognitive load (Reber & Kotovsky, 1997).

This finding would seem to indicate that implicit learning is just as susceptible to working memory limits as explicit learning. Assuming that the cognitive load manipulation demanded attentional resources, such a conclusion would be generally consistent with Nissen's (1992) finding that attention is required for implicit learning to occur. However, it is also possible that the task presented performance goals that engaged both explicit and implicit learning processes. Participants in this study had an explicit goal of solving the puzzle and were presumably utilizing explicit hypothesis testing processes in order to learn the problem solution initially. This alone would explain the effects of working memory load on the task. Consistent with this, when solving the problem for the second time, and on subsequent trials, the working memory load had no effect on the participants solving the problems, as those participants had equivalent times to those not under a working memory load (Reber & Kotovsky, 1997).

In sum, there is evidence that suggests working memory limitations may be detrimental to implicit memory processes in addition to explicit processes. However, explicit processing demands may confound this evidence, and there is other evidence that implicit memory processes may be less impacted by the limits of working memory (Warmington, Hitch, & Gathercole, 2013; Woltz, 1988). Given the existing evidence, if the goal of implicit learning methods is to reduce working memory demands, it would be important to structure the learning task so as to avoid attention being drawn to irrelevant task features and to minimize conditions likely to engage explicit memory processes.

Errorless Learning and Vanishing Cues

In a study using motor learning of a golf-putting task, Maxwell, Masters, Kerr, and Weedon (2001) investigated procedural learning by having participants practice putting through an errorless learning paradigm. The purpose of this errorless learning was for participants develop putting skill without any declarative instructions while making few, if any, mistakes. This was accomplished by initially positioning participants very near the target hole and gradually moving them further away over subsequent trials. This allowed participants to make very few errors while also receiving no instructions. The study also included a control group that learned to putt in a hypothesis-testing paradigm that was explicit in nature. The group using the errorless implicit learning paradigm experienced robust skill performance that did not degrade under stress or attentional demands. In contrast, the performance of those who learned using the explicit method was degraded under stress and attentional demands.

In a study that continued the utilization of the errorless putting paradigm, Poolton, Masters, and Maxwell (2005) contrasted this form of errorless learning with a condition that utilized initial explicit instructions. One group began with the errorless procedural learning of the putting task and only received declarative instructions later in the task. The other group received the declarative instructions before any practice and was then able to practice putting. The group that began with errorless learning performed significantly better than the instructions-first group when a secondary task load was introduced. The errorless learning group also performed significantly better in transfer compared to the instructions first group and actually showed no degradation in the transfer task. The authors concluded "the possibility of retaining the advantages of a consciously accessible knowledge base while offsetting the negative consequences of explicit learning via the insertion of an initial period of implicit learning provides a practical alternative to previous solutions" (Poolton et al., 2005, p. 376). This evidence suggests that the use of the errorless learning paradigm in place of initial instruction can minimize the working memory load during initial task learning, apparently allowing for better skill acquisition and transfer.

There have also been a variety of studies investigating errorless learning in more academic domains, particularly in various word learning tasks; in general these tasks have confirmed the benefit of errorless learning (Anderson & Craik, 2006; Baddeley & Wilson, 1994; Hunkin, Squires, Parkin, & Tidy, 1998; Page, Wilson, Shiel, Carter, & Norris, 2006; Tailby & Haslam, 2003; Warmington et al., 2013). While the majority of these studies were conducted using adults, both healthy and impaired, there have also been studies confirming the benefits of errorless learning in children (Warmington et al., 2013).

Many of these studies sought to investigate the mechanisms that drive the benefit found in errorless learning. In their study, Hunkin and colleagues (1998) reported that the benefits of errorless learning were not the result of implicit mechanisms of memory, but instead stemmed from error correction mechanisms in residual explicit memory. There have been several studies since that provided evidence to the contrary. In fact, according to Anderson and Craik, (2006) "errorless learning works through implicit means" (p. 2811). They, like others, found that errorless learning was likely an implicit process.

A recent study of children provided more evidence for this idea; in the study, children learned to associate nonwords with novel images in either errorful or errorless learning conditions. The errorless condition consisted of presenting children with an image and the first letter of its associated word, closely followed by the word itself; children then recorded the word; in contrast, children in the errorful condition were given the first letter and required to guess the word, which was presented if not guessed (Warmington et al., 2013). Warmington and colleagues (2013) not only found errorless learning to be beneficial in children, but they also suggested that the "independence of errorless learning from cognitive skills known to be important in explicit memory…arises because it relies instead on implicit memory." (p. 462).

Tailby and Haslam (2003) conducted a study investigating the effects of self-generation on errorless learning; memory impaired participants learned lists of words in one of three learning conditions: errorful, standard errorless, or modified errorless featuring self-generation of responses. The self-generation in the modified errorless condition consisted of participants being presented with the first two letters of a word, its length, and many contextual clues as to its identity; participants were then instructed to generate the target word and record it once correctly identified. In contrast, those in the standard errorless condition were presented the first two letters of a word and its length and were then given the word and instructed to record it. Those in the errorful condition were presented the first two letters and instructed to generate the word by guessing. Participants in the modified errorless condition outperformed those in the standard errorless condition during the target word recall post assessment. In addition, the standard errorless group outperformed the errorful group by a similar margin, seemingly indicating a benefit of self-generation nearly equal to the independent benefit of errorless learning over errorful learning.

The method of vanishing cues is also used in Implicit Learning Tasks when the goal is to provide relatively error free practice without initial knowledge of declarative rules. The method of vanishing cues dictates that participants in a task are given cues for correctly answering a question or problem and that the cues are gradually diminished over the course of practice. This method has been used in learning for individuals with declarative memory impairments because it relies on implicit rather than explicit learning (Evans, Levine, & Bateman, 2004; Riley, Sotiriou, & Jaspal, 2004). However, the method of vanishing cues has had somewhat mixed results, as some research has shown no learning effects (Kessels & de Haan, 2003). It is suggested that the lack of significant effects of the vanishing cues in that research may be related to participants making mistakes. This could potentially be alleviated in a learning task that combines other errorless learning procedures and vanishing cues.

Worked Examples

The worked examples approach involves showing learners mathematics problems that have been solved, with all the 'work' shown along with the solution in order to improve learning (Atkinson, Derry, Renkl & Wortham, 2000; Renkl & Atkinson, 2007; Renkl, Atkinson, & Grosse, 2004). The worked examples approach to math learning is similar to some forms of implicit learning in that providing worked examples likely provides some form of implicit learning of solution patterns through exposure to correct answers. In addition, a strand of this research has found that gradual fading of worked examples steps fosters increased skill acquisition (Renkl et al., 2004), which is consistent with principles of error free learning (Evans et al., 2004; Maxwell et al., 2001; Riley et al., 2004) and suggests increased likelihood that error free learning can be applied in the context of math learning. Although the worked examples research shares some elements with the idea of implicit learning, it is important to recognize that in the worked examples approach "the basic domain principles are typically introduced by a text" (Renkl & Atkinson, 2007); in this way, the worked examples approach is consistent with Anderson's (1982, 1993) theories. The worked examples approach also differs from the implicit methods in that study participants are expected to explicitly understand the solution steps provided in the worked examples through self-explanation activities (Renkl, Atkinson & Grobe, 2007) and making analogies between worked examples and problems to be solved (Atkinson et. al., 2000, p. 185).

Order of Learning

There have been studies on the order of instructional tasks and its impact on learning. In a study by Schwartz and Bransford (1998), college students prepared for a lecture or course reading by analyzing relevant contrasting cases prior to the learning event. According to the authors, analyzing these contrasting cases before attending the lecture or reading the text resulted in a better quality of learning, as evidenced by a prediction task, than those who did not contrast cases beforehand (Schwartz & Bransford, 1998). The analysis task performed before the lecture can be described as exposure to a problem domain prior to declarative instruction in that domain and the enhanced learning resulting from it may be applicable to implicit learning of problem patterns in mathematics.

A study by DeCaro and Rittle-Johnson (2012) investigated the use of exploration before explicit instruction to increase math learning. Their experiment centered around elementary aged students solving simple addition equivalence statements, like 3 + 7 = 4 + [ ], either before or after explicit instruction on the concept of mathematical equivalence and the equal sign; in each case students were given accuracy feedback after solving each statement. In each condition, some students solved additional problems, and some students were given self-explanation prompts after each item, but all students were able to use pencil and paper and were asked to report their solution strategy after each problem. Procedural and conceptual knowledge were measured with separate posttests after the interventions. While there were no effects of condition on procedural knowledge of solving problems, students who solved problems before instruction performed better on the conceptual knowledge test. There were also no differences between those who solved additional problems and those who self-explained (DeCaro & Rittle-Johnson, 2012, p. 560), suggesting that declarative processing of problem solving strategies did not enhance the value gained from skill practice.

The results of this study are consistent with the work of Schwartz and Bransford (1998) and suggest that actively working with information in a domain supports later learning in that domain. Despite the declarative nature of the problem-solving task (participants reported their problem solving strategies after each problem) in DeCaro and Rittle-Johnson's (2012) study, the evidence supporting the value of exposure to problem structure before knowledge or rule learning is encouraging. Furthermore, in both of these studies it is feasible that the participants may have also gained some concurrent implicit understanding of patterns present in the problems.

Math Instruction

Math Curricula

In a document reflective of current practice in mathematics education, the National Council of Teachers of Mathematics (2000) published the *Principles and Standards for School Mathematics*, which laid out best practices for teaching mathematics in addition to standards of math concepts to be taught from kindergarten through 12th grade. This document described a variety of mathematic teaching methods including discussions, manipulative work, and thought provoking questions. While varied, these methods share a common thread in the explicit nature of their instruction for problem solving and algorithm dependent procedures.

These methods are used to help students make connections with challenging declarative knowledge for understanding procedural mathematical skills. Manipulative work, for example, is intended to help students by providing another way to represent math problems. This representation is intended to be concrete in nature in order for children to more easily interpret it. The rationale is that young children can more easily comprehend abstract math concepts, like numbers, by being able to see a physical representation of the abstract concept, for instance, using cubes to represent numbers when children are learning to add. While there certainly may be a perceptual component to working with manipulatives, the goal is to help children gain a declarative representation of the given concepts; manipulatives could easily be described as a different symbol for representing problems. Discussions and questions have similar goals, in that they are intended to help students think about challenging mathematics concepts and skills in different, declarative ways. All of these methods seek to find ways for students to think about math that are easier for them to understand, but they rely on working memory to comprehend, remember and apply verbal explanations for solving math problems.

Math textbooks

A review of many mathematics textbooks for a variety of grade levels and from various publishers provides more information about common methodology for teaching mathematics skills and concepts. Saxon math is one such textbook series that relies heavily on declarative mathematics instruction. A review of the *Algebra 1* (Saxon, 2003) text revealed a typical method for teaching algebra problem solving skills and concepts; new problem solving skills are initially introduced through verbal descriptions of rules and methods. Figure 1 shows the rules presented in this textbook for factoring a trinomial, while Figure 2 shows the related problem solving method of applying the rules in an example problem. This common method of mathematics instruction relies on students understanding declarative rules and then applying them to problems.

In addition to the Saxon math book, I also reviewed several textbooks written for students ranging from 3rd grade–high school from a variety of publishers. While the methods and techniques for teaching math skills vary among these textbooks, they share a common reliance on the initial use of declarative instruction to teach, and more specifically introduce, new concepts and skills (Bell, 1998; Bumby, Klutch, Collins, & Egbers, 1995; Charles, Branch-Boyd, Illingworth, Mills, & Reeves, 2004a, 2004b; Lappan, Fey, Fitzgerald, Friel, & Phillips, 1998; Larson, Boswell, Kanold, & Stiff, 2007; Leschensky, Malloy, Price, Rath, & Alban, 1999). This is consistent with the Saxon text, and, moreover, I believe these texts to be representative of the textbooks used in public and private schools throughout the United States.

Math Achievement

A review of *The Nation's Report Card: Mathematics 2011* (National Center for Education, 2011) presents statistical information that can aide in the evaluation of current math instruction methods. This document presents student performance data from the "National Assessment of Educational Progress (NAEP), a continuing and nationally representative measure of achievement" (National Center for Education, 2011, p. ii). Measures of achievement for grades 4 and 8 are reported, and student performance is reported as having met criteria for three achievement levels of math proficiency: basic, proficient, and advanced. The basic level "denotes partial mastery of prerequisite knowledge and skills that are fundamental for proficient work at each grade" (National Center for Education, 2011, p. 7). The proficient level indicates a student has mastered the fundamental concepts at grade level and is the goal level for student achievement. The advanced level indicates superior math performance. Students who do not meet the criteria for the basic level are considered to have below a basic level of math skills.

While trends over the last 10 years have shown slow growth in achievement levels of 4th and 8th grade student in math achievement, results of the 2011 assessments indicate there is still much room for improvement. On the 2011 NAEP math assessment, 60% of 4th grade students were below the proficient level of achievement, and 18% were below the basic level of math achievement (National Center for Education, 2011, p. 2). Eighth grade students performed worse, with 65% of student achieving below the proficient level. In addition, 27% of 8th grade students were below the basic level of achievement. These results indicate that nearly a fifth of 4th graders and more than a fourth of 8th graders lack even a basic understanding of fundamental math skills for their respective grades. Beyond this, less than half of the nation's 4th and 8th grade students are proficient in their use of fundamental math skills.

Given the current standards and curricula in the area of mathematics instruction, the prevalent methods for teaching math skills and concepts emphasize the declarative understanding of rules and algorithms. Overall student achievement levels in mathematics are disappointing from this form of conventional curriculum, although reasons for the low achievement nationally certainly go beyond the curriculum and traditional methods of instruction. Nevertheless, the body of laboratory research in the area of implicit procedural learning begs the question of whether those ideas and principles can be applied in the context of mathematics skill learning. Learning complex algorithms and problems solving skills in mathematics is often a difficult and frustrating process for children. Ashcraft and Krause (2007) suggest that, "Math is ... a cognitively challenging topic… the stage is set early on in math education for students to be 'stranded' without a reasonable, instructive explanation for many aspects of math" (pp. 246–247). It is feasible that working memory limitations may be part of what makes the acquisition and application of these complex algorithms so difficult for many children. In fact, Ashcraft and Krause state, "There is a pervasive reliance throughout arithmetic and math on the working memory system, from simple counting and estimation processes up through algebra and complex problem solving" (Ashcraft & Krause, 2007, p. 246). Given that emphasis on initial declarative learning of these challenging skills might unduly tax working memory, it would be of value to explore the initial use of implicit learning methods to decrease working memory demands in mathematics instruction.

Research Questions

Question One

Can algebraic problem solving skill be acquired without exposure to the declarative rules? This research seeks to determine whether it is possible to implicitly obtain procedural knowledge of the patterns of solutions for factoring algebraic polynomials. The implicit training to evaluate this question provides extensive exposure to problem and solution patterns for factoring algebraic polynomials with no declarative training. Acquisition of procedural knowledge through this training will be evaluated with measures of problem solving administered prior to declarative instruction. I predict participants engaging in initial implicit training will be able to acquire problem-solving skills in the absence of declarative training.

Question Two

Does initial exposure to an implicit learning condition for algebraic problem and solution patterns facilitate the subsequent learning of declarative rules for solving those problems? I predict that participants engaging in implicit-first training will perform better on measures of declarative rule learning than those participants who engaged in initial declarative training based on standard textbook instruction. This effect is predicted because an implicit recognition of pattern solutions for various problem types should reduce the working memory demands associated with encoding and analyzing verbal descriptions of the problem components.

Question Three

Does initial exposure to an implicit learning condition for algebraic problem and solution patterns have an effect on student perceptions of the difficulty of learning in this domain? It is valuable to know whether students feel the process is easier as a result of initial implicit training, as mathematics anxiety can be a major factor in how much mathematics students pursue. It is predicted that exposure to initial implicit training will result in reduced participant difficulty perceptions of problem solving and rule learning, as compared to those who engaged in initial declarative training. Again, this effect is predicted because of a presumed reduction in working memory demands following the acquisition of implicit knowledge about problem patterns.

Question Four

Does initial implicit learning prior to declarative instruction of algebraic problem and solution patterns result in better final problem solving and transfer performance compared to learning declarative rules prior to the procedural practice? This question contrasts the effectiveness of a more traditional sequence of math instruction (learning declarative knowledge about problem solving before practice) with the reverse order of instruction that is designed to avoid initially high working memory demands that presumably impede learning in many students. It is predicted that participants in the reverse ordered, implicit-first learning condition will outperform those in the declarative-first training on final measures of problem solving and transfer performance.

- The first term of the trinomial is the product of the first terms of the binomials
- The last term of the trinomial is the product of the last terms of the binomials
- The coefficient of the middle term of the trinomial is the sum of the last terms of the binomials
- If all signs in the trinomial are positive, all signs in both binomials are positive. If a negative sign appears in the trinomial, at least one of the terms of the binomials is negative.

Figure 1. An example of rules for factoring a trinomial from *Algebra 1*.

We use these observations to help us factor trinomials. To factor the trinomial

$$x^2 - 3x - 18$$

we first write down two sets of parentheses to form an indicated product.

$$(\quad)(\quad)$$

Since the first term in the trinomial is the product of the first terms of the binomials, we enter x as the first term of each binomial.

$$(x \quad)(x \quad)$$

Now the product of the last terms of the binomials must equal -18, their sum must equal -3, and at least one of them must be negative. There are six pairs of integral factor of -18:

$(-18)(1) = -18$	$(2)(-9) = -18$	$(3)(-6) = -18$
$(18)(-1) = -18$	$(-2)(9) = -18$	$(-3)(6) = -18$

Their sums are

$(-18) + (1) = -17$	$(2) + (-9) = -7$	$(3) + (-6) = -3$
$(18) + (-1) = 17$	$(-2) + (9) = 7$	$(-3) + (6) = 3$

Note that while all six pairs have a product of -18, only one pair (3 and -6) sums to -3. Therefore, the last terms of the binomials are 3 and -6, and so $(x + 3)$ and $(x - 6)$ are the factors of $x^2 - 3x - 18$ because

$$(x + 3)(x - 6) = x^2 - 3x - 18$$

Figure 2. The method for teaching polynomial factorization using an example problem modified from *Algebra 1* (Saxon, 2003, p. 281).

CHAPTER 3

METHOD

Participants and Apparatus

Participants were students enrolled in an introductory algebra course from two different southeastern Minnesota schools in the same school district and city. The algebra course was intended to be equivalent across schools in the district and both schools utilized the same textbook; the course was also a prerequisite for a high school algebra course. The participants were volunteers who received no compensation or course credit for participation, although participation did occur during the normal algebra class time. Of the original 188 participants, 17 (9.0%) were eliminated as outliers based on their performance on daily learning tests (described later). The final sample (N = 171) included 103 females and 68 males all in 8^{th} grade. An additional 19 participants took part in training that was a supplemental condition added to evaluate the accuracy of one of the measures used.

All participants performed both declarative and Implicit Learning Tasks and assessments of learning on computers in the school computer lab. The lab consisted of Windows based computers using a standard display and keyboard. The tasks were controlled by E-Prime 2.0 runtime software (Schneider, Eschman, & Zuccolotto, 2002). The program was created using the E-Studio software from E-Prime 2.0. Participant responses were made using specified keys on the keyboard corresponding to response alternatives shown on the computer display. Instructions for all components of the experiment were presented over headphones or by text on the computer display.

Design and Procedure

The experiment consisted of three sessions over 3 days, with participants randomly assigned to either the implicit learning first condition or the declarative learning first condition. Participants from one of the schools (N = 90) completed the three sessions on three consecutive days, while the participants from the other school (N = 81) experienced a 4 day gap between Sessions 2 and 3 because of a weather related school closing. Table 1 summarizes the sequence of tasks in the two experimental conditions. All participants completed a knowledge pretest at the beginning of Session 1. Students within each classroom were randomly assigned to one of the two learning conditions. On both the first and second days of the experiment participants in each group were exposed to the initial learning condition to which they were assigned (implicit or declarative), and they took the learning test at the end of each session. On the third day, participants completed the learning task they did not perform on the first 2 days. They finished the session by taking the final learning test and the transfer tests. An attempt was made to equate the three learning tests for difficulty, and assignment to session was randomized so that learning tests administered during each session were assumed to be equivalent on average. Both groups completed a declarative learning test before and after the first (or only) session of declarative learning. This occurred during Session 1 for the declarative learning first group and on Session 3 for the implicit learning first group.

Table 2 summarizes the polynomial forms and sign patterns presented to participants throughout the study. For each of the two polynomial forms, $x^2 + Bx + C$ and $x^2 + Bx$, there were four different sign patterns that yielded different solution patterns. For each polynomial form, two sign patterns were used in the learning tasks and on the learning tests; the other two patterns were reserved for the near transfer test. These sign patterns are shown in the Learning and Near Transfer section of Table 2. In order to avoid a potential confound, the sign patterns for each polynomial form were counterbalanced, such that half of the participants saw the near transfer patterns of Table 2 in learning and the learning patterns on the near transfer test.

Declarative Learning Task

Participants listened to oral directions explaining how to factor polynomials while viewing cued visual images of polynomials and solutions relevant to the directions being presented. Participants were instructed in the rules for factoring polynomials of the forms $x^2 + Bx + C$ and $x^2 + Bx$ and were able to progress through the instruction frames self-paced. Table 3 shows the structure of the Declarative Learning Task. Upon completion of a slide, participants were able to move on to the next slide, repeat the slide including the audio component, or go back to the previous slide by using the right, down, and left arrows, respectively.

The rules and content of this learning task were adapted directly from the textbook *Algebra 1: An Incremental Development* (Saxon, 2003). The only major adaptations to the content were that the textual information was presented in an audio format alongside the visual elements from the text, and visual cues (e.g., arrows indicating relevant content) were used to maximize the connection between visual and aural information. This adaptation was made based on multimedia learning research that indicates the best learning outcomes are achieved through aurally presented textual information alongside visual elements (Mayer & Moreno, 1998; Moreno & Mayer, 1999). In addition, at the onset of the first (or only for participants in the implicit-first condition) session of declarative learning, participants were presented with slides introducing and defining relevant terms and the parts of a polynomial needed to learn the declarative rules. The second session for the participants in the declarative-first condition was identical to the first session, but without the terminology introduction.

Participants were initially introduced to the rules, aurally and visually, for factoring a given type of polynomial. These were the rules presented in Figure 1 in Chapter 2. After being introduced to the rules for each polynomial type, participants were exposed to two example problems, in which aural descriptions explained how to apply the rules to factor the polynomial. Table 2 shows the exact example problems used; due to the counterbalancing, half of the participants saw the problems in the Learning column and half saw the problems in the Near Transfer column. Figure 3 shows the sequence of slides for one example problem in the Declarative Learning Task. Figure 3 parts a, c, e, g, i, and k (on the left) are the visual slides seen by the participants, while parts b, d, f, h, j, and l (on the right) are transcripts of the audio that participants heard for each slide. The arrow cues in these figures were synchronized with the aural descriptions to connect the aural and visual information.

Implicit Learning Task

Participants factored polynomials of the same forms present in the Declarative Learning Task by selecting the correct factorization from two choices. Figure 4 is an example of the slide format used in the Implicit Learning Task. They were given no directions about how to solve the problems, only to select an option by pressing either the C or M key. Participants completed 9 blocks of 24 items of this type per session of implicit learning. Each block contained 12 polynomials of each of the two forms, further divided into two sets of six polynomials with the same sign pattern. During Session 1 of this task, which participants from both conditions completed albeit on different days, the patterns were presented in a sequential format. All six items of a given sign pattern and polynomial form were presented in sequence; this was repeated for each of the other three patterns in each block, but the order of patterns presented within a block was randomized. Table 2 details the polynomial forms and sign patterns seen during this task. Session 2 of the Implicit Learning Task was completed only by the participants in the implicit-first learning condition. The polynomial patterns in this session were presented in an alternating format, such that each group of four items included all four polynomial and sign patterns. Equivalent, randomly assigned sets of numbers were used to create problems for the two sessions, so the problems were not identical.

The 9 blocks of a session were separated into sets of three blocks with different types of foils. The first set of 3 blocks had foils that represented completely different types of polynomials. The second set of 3 blocks had foils from polynomials of the same form, but with different patterns of positive and negative numbers. The third set of 3 blocks had foils from exactly the same polynomial type, with only the actual numbers distinguishing the foils from the correct answer. Figure 5 illustrates the types of foils used in each of the 3 sets of blocks; only 1 foil appeared in any trial slide

Within each set of 3 blocks, the first block had numbers in the foils that were distant from the numbers in the correct answer, in that they produced neither the correct product nor sum to accurately factor the polynomial. The second block contained numbers that either produced the correct sum or the correct product for an accurate polynomial factorization. The third block contained the same numbers as the correct solution, except in the third set, which used the second block number pattern in order to avoid having two correct answers.

This patterning of foils throughout the implicit learning blocks is built on the idea of vanishing cues, in that the discriminability of the incorrect and correct answer choices slowly decreased over the course of the 9 blocks. Stated otherwise, the foils initially appeared very different from the correct answers and became nearly identical to the correct answers by the final blocks. This format was intended to minimize mistakes while providing participants with an opportunity to implicitly learn the correct answer patterns. In addition, it required finer discriminations between correct answers and foils as the blocks progressed in order to promote detailed rather than superficial pattern recognition.

When participants selected an answer choice, they were given feedback as to the accuracy of their selection. Figures 6 and 7 are examples of implicit learning slides after correct and incorrect answers are recorded, respectively. When the correct answer was selected, the word "correct" was displayed in the center of the screen. When the incorrect answer was selected, a large red "X" appeared over the incorrect answer, obscuring it, and a green box enclosed the correct answer. In both cases the feedback was presented for a fixed amount of time for each block, with the amount of time systematically decreasing from block one (1500 ms) to block nine (750 ms). In this way, the correct factorization of a polynomial was highlighted whenever a participant made an error. The incorrect feedback was presented throughout the nine blocks of items, but the correct feedback faded from a dark color in the first blocks, to a light color in the middle blocks, to no correct feedback for the last few blocks. This gradually fading feedback style was intended to highlight the connection between a polynomial and its correct factorization, without giving any explicit explanation.

The Implicit Learning Task described here allowed participants to get extensive amounts of practice factoring polynomials without being given any declarative rules or directions about how to do so. The combination of simple, language free accuracy feedback and vanishing cue style foil pattern was intended to minimize errors, strengthening the implicit recognition of the factorization pattern for each type of polynomial.

Learning Tests

A learning test was taken immediately after both the implicit and Declarative Learning Tasks and was intended as an assessment of students' understanding of how to factor polynomials. Participants completed 24 items on each learning test, using the same polynomial forms and sign patterns as in the learning tasks, but the specific items were different than those seen during either of the learning tasks. In this task, participants were presented with a polynomial in the center of the screen as well as a partial factorization that was missing one of the numbers as shown in Figure 8. Participants entered the correct number using the numeral keys to complete the factorization. No feedback was provided. The learning tests were designed with the intent of assessing knowledge of polynomial factoring without favoring either learning condition. The requirement to generate a numeric response presumably required a different problem solving process than that used in the errorless learning format of the implicit task. Although the declarative-first condition did not provide any problem solving practice, determining a missing number from one binomial in the solution required a straightforward application of a portion of the rules that had been presented multiple times along with example problems.

Near and Far Transfer Tests

The transfer tests were formatted identically to the learning tests. Participants entered numbers to complete a partial factorization, and they received no feedback. However, the transfer tests' content represented both near and far transfer of learned skills. Participants completed 24 items on the near transfer test and 36 items on the far transfer test in separate test sections. The near transfer items consisted of polynomials of the same two forms presented in the learning tasks, but with two patterns of negative and positive numbers that were not presented in learning for each of the two polynomial forms (see Table 2). The far transfer items consisted of polynomial forms not seen by the participants in learning: $x^2 – C$ and $Bx – C$. The polynomial forms $x^2 – C$ and $Bx + C$ are factored similarly to their counterparts in the learning tasks,

$x^2 + Bx + C$ and $x^2 + Bx$, respectively, but have features that make the application of the factorization different from the polynomials learned earlier (See Figures 9 and 10).

Declarative Rules Test

The declarative rules test was designed to assess how well participants learned the verbal rules for solving polynomials. It consisted of an initial section in which participants generated the rules for each polynomial type, and a subsequent section in which participants identified the rules that are used to factor a given type of polynomial. In the first section participants were presented with each of two polynomial forms, with rectangles in place of numbers, and instructed to use the keyboard to type the rules for factoring this polynomial. There were no time or character limits for this section of the test. In the latter section, participants were presented with an example of a polynomial and were required to identify which of several presented rules applied to the given polynomial. Figure 11 is an example of a slide from this section of the declarative rules test. Participants completed this test as a pretest and a posttest administered immediately before and after their initial Declarative Learning Task. There were ten items on the test, five for each polynomial form. There were three answer options for each item; two options were rules and the third indicated that both rules were correct for that polynomial form. The pretest and posttest presented the same rules for polynomial factorization, but different items were constructed for each test by pairing different rules for each item.

There was also a modified version of this pretest administered to a small independent sample of participants who performed the implicit learning condition. These participants were asked at the beginning of the pretest to explain how they would teach this skill to a peer; this condition was otherwise identical to the condition in which participants generated rules for simplifying polynomials.

Knowledge Pretest

Participants took a paper pretest before beginning the experiment in which they attempted to factor polynomials of the types that were present in the experiment. There were 12 items on the pretest representing the polynomial forms and sign patterns presented in the learning, near transfer, and far transfer sections of the study. Polynomials were displayed on a sheet of paper with directions to factor each polynomial into a product of binomials. The data from this pretest were designed to exclude participants who already knew how to factor polynomials.

Difficulty Questionnaire

Participants completed a questionnaire consisting of two items at the end of Session 3. The questionnaire was intended to assess participants' perceptions of the difficulty of solving the problems and learning the rules. The items required participants to respond to statements about learning difficulty using a 5-point scale. Figures 12 and 13 show the Difficulty Questionnaire items for rule learning difficulty and problem solving difficulty, respectively.

Table 1. Experimental design.

	Implicit Learning First Group	**Declarative Learning First Group**
Pre Session	• Knowledge PreTest	• Knowledge PreTest
Session 1	• Implicit Learning Task • Fill-in-the-Blank Learning Test	• Declarative Pretest • Declarative Learning Task • Declarative Post-Test • Fill-in-the-Blank Learning Test
Session 2	• Implicit Learning Task • Fill-in-the-Blank Learning Test	• Declarative Learning Task • Fill-in-the-Blank Learning Test
Session 3	• Declarative Pretest • Declarative Learning Task • Declarative Post-Test • Fill-in-the-Blank Learning Test • Fill-in-the-Blank Near Transfer Test • Fill-in-the-Blank Far Transfer Test • Difficulty Questionnaire	• Implicit Learning Task • Fill-in-the-Blank Learning Test • Fill-in-the-Blank Near Transfer Test • Fill-in-the-Blank Far Transfer Test • Difficulty Questionnaire

Table 2. Example of polynomial forms and sign patterns.

Polynomial Form	Learning[a]	Near Transfer[b]	Far Transfer
$x^2 + Bx + C$	$x^2 + 9x + 18$ $x^2 – 6x – 16$	$x^2 – 9x + 18$ $x^2 + 6x – 16$	$x^2 – 16$
$x^2 + Bx$	$x^2 – 5x$ $-x^2 + 8x$	$x^2 + 5x$ $-x^2 – 8x$	$4x + 12$ $4x – 12$ $-4x + 12$ $-4x – 12$

[a] Half of the participants saw these patterns as near transfer items
[b] Half of the participants saw these patterns during learning

Table 3. Declarative Learning Task structure.

1st Session	• ***Explanation of Relevant Terminology & Symbols***
	    o *Polynomial, binomial, trinomial*
	    o *Parts of polynomials*
	• ***x^2 + Bx + C Pattern***
	    o Rule presentation
	    o Example 1 presentation
	    o Example 2 presentation
	• ***x^2 + Bx Pattern***
	    o Rule presentation
	    o Example 1 presentation
	    o Example 2 presentation

2nd Session	• ***x^2 + Bx + C Pattern***
	    o Rule presentation
	    o Example 1 presentation
	    o Example 2 presentation
	• ***x^2 + Bx Pattern***
	    o Rule presentation
	    o Example 1 presentation
	    o Example 2 presentation

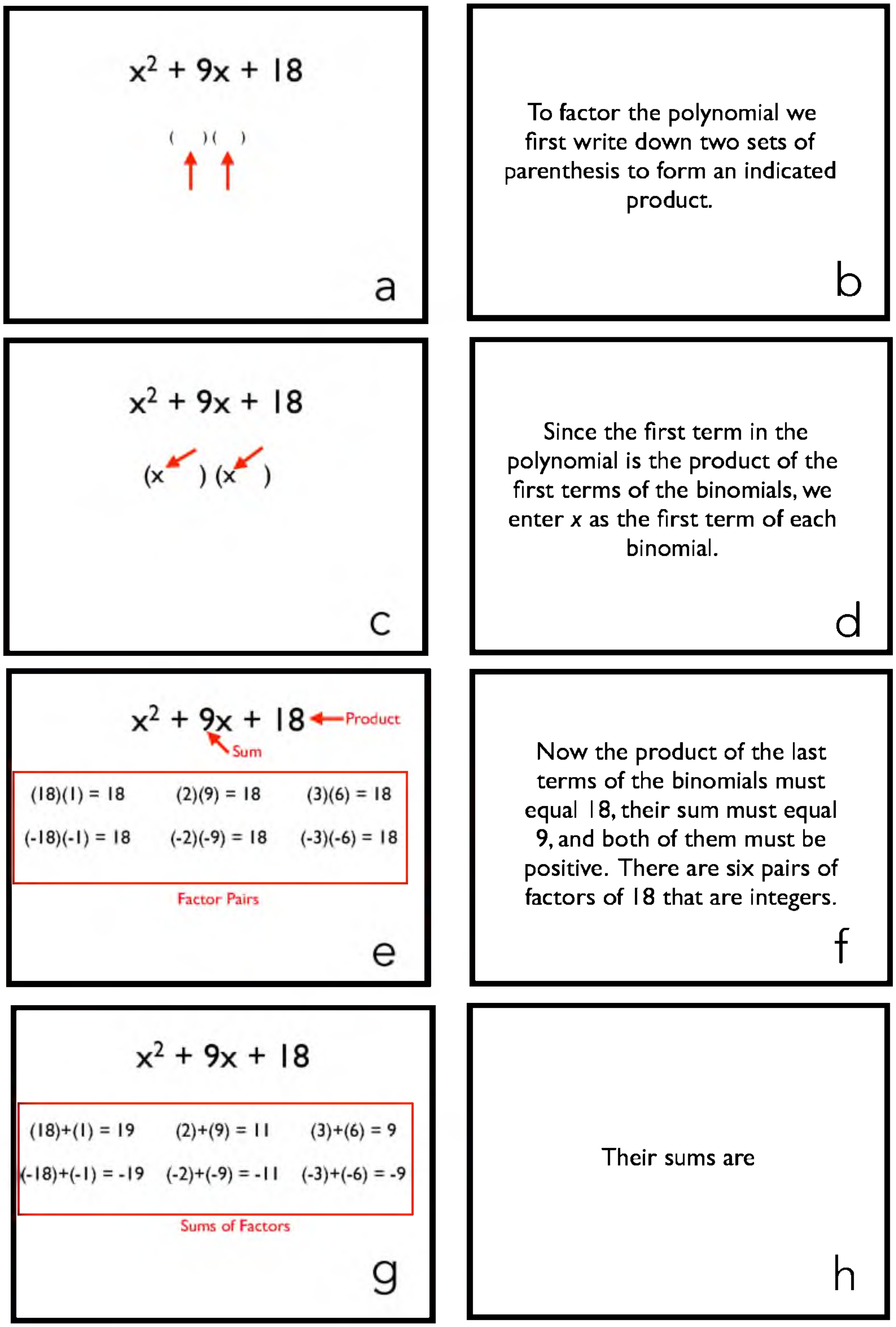


Figure 3. Slides and audio transcript of a worked example in the Declarative Learning Task.

a, c, e, g, i, and k are visual slides

b, d, f, h, j, and l are audio transcripts heard during slides a, c, e, g, i, and k, respectively.

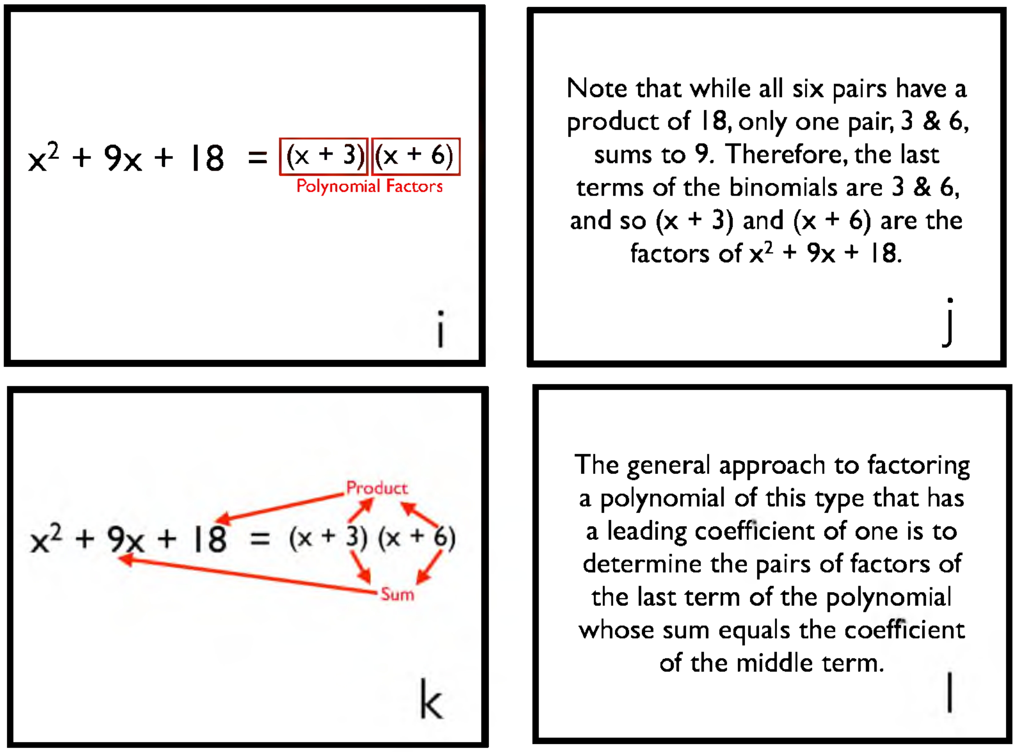


Figure 3 continued.

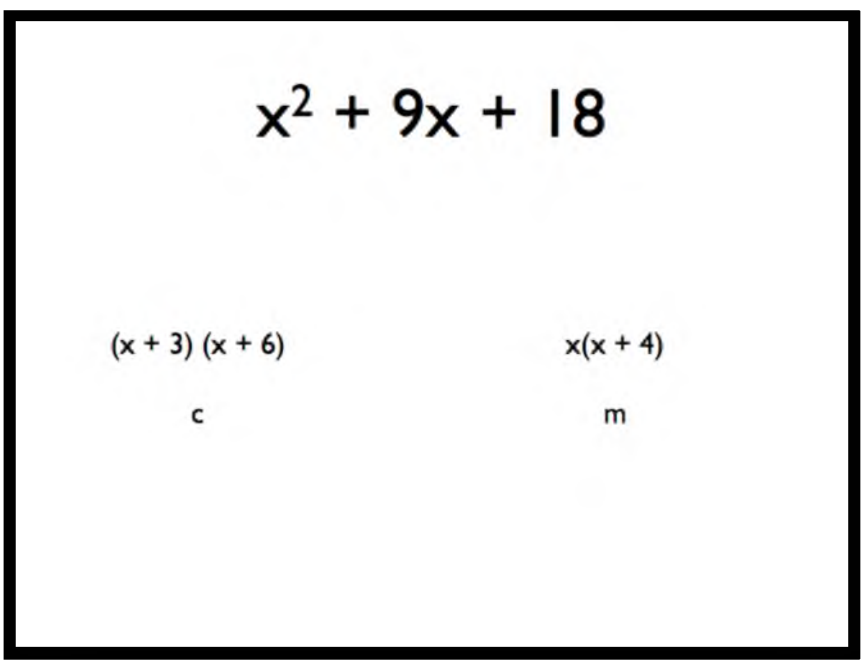


Figure 4. Example slide from the Implicit Learning Task.

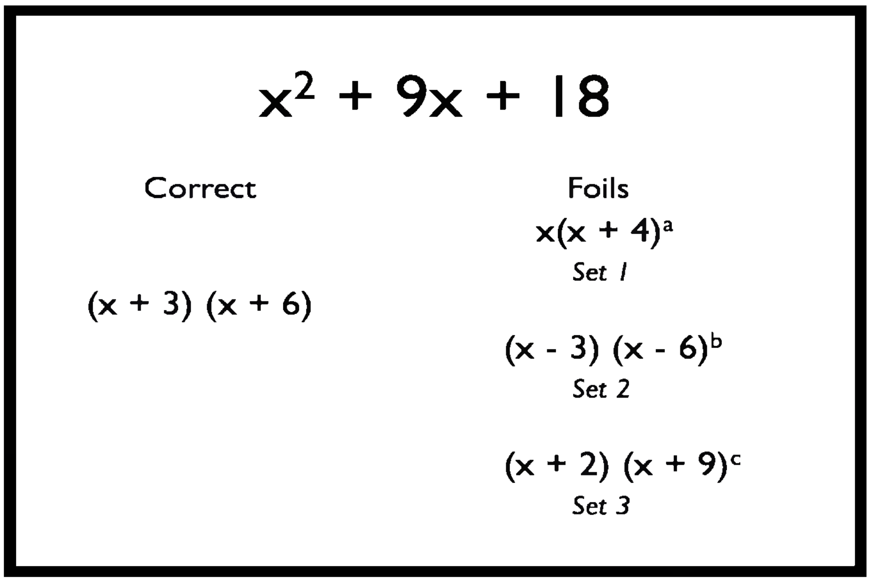


Figure 5. Trial slide with examples of foils.

[a] This foil is seen in Set 1, blocks 1–3, and is from a polynomial of the form $x^2 + Bx$.

[b] This foil is seen in Set 2, blocks 4–6, and is from a polynomial of the form $x^2 – Bx + C$.

[c] This foil is seen in Set 3, blocks 7–9, and is from a polynomial of the form $x^2 + Bx + C$.

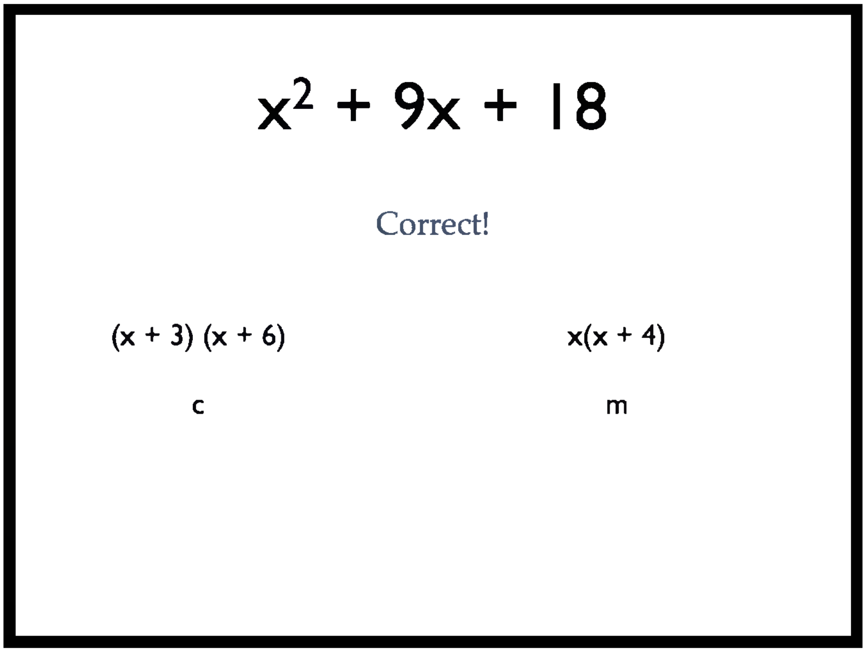


Figure 6. Correct response feedback.

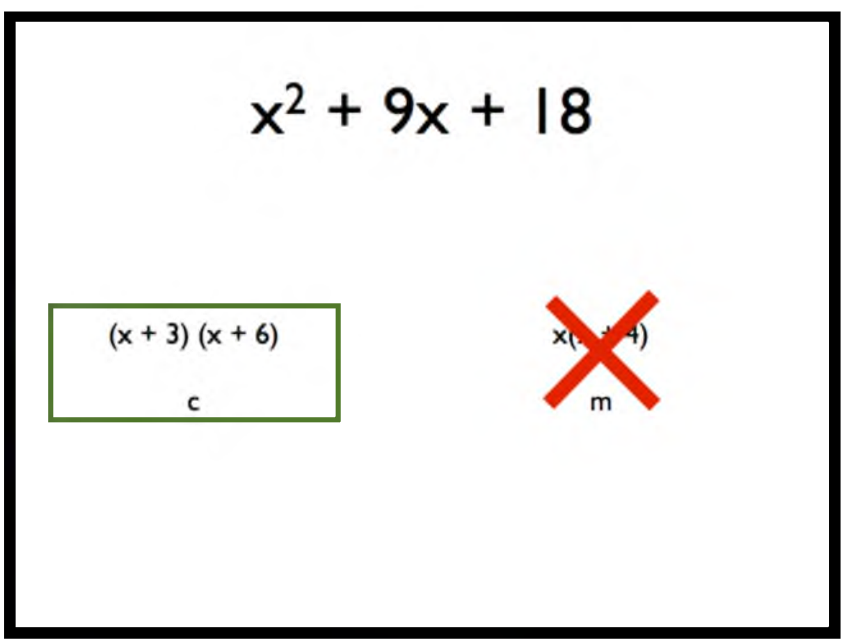


Figure 7. Incorrect response feedback.

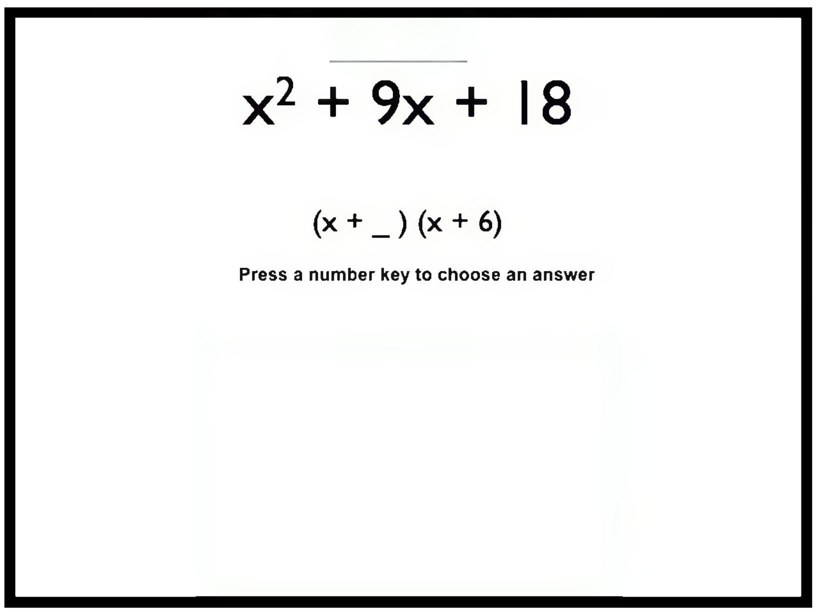


Figure 8. Example slide from the learning test.

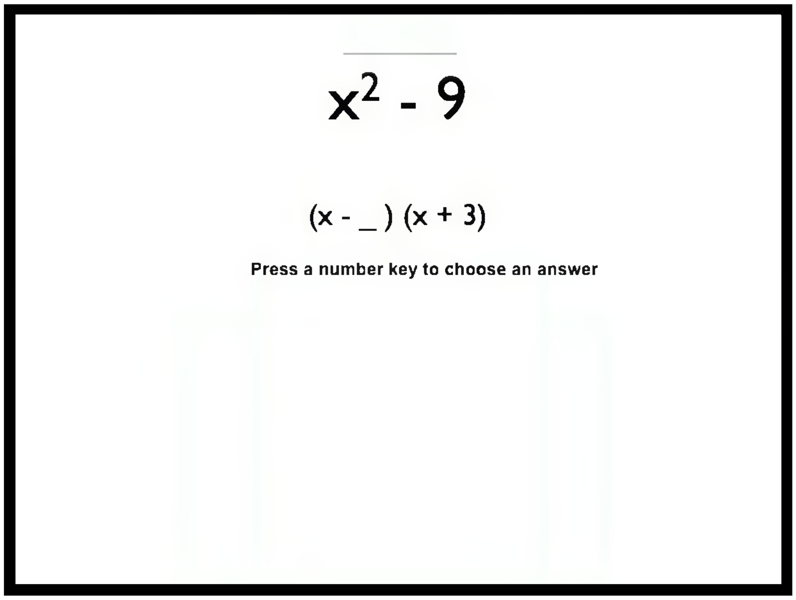


Figure 9. Example far transfer item slide: polynomial form $x^2 - C$.

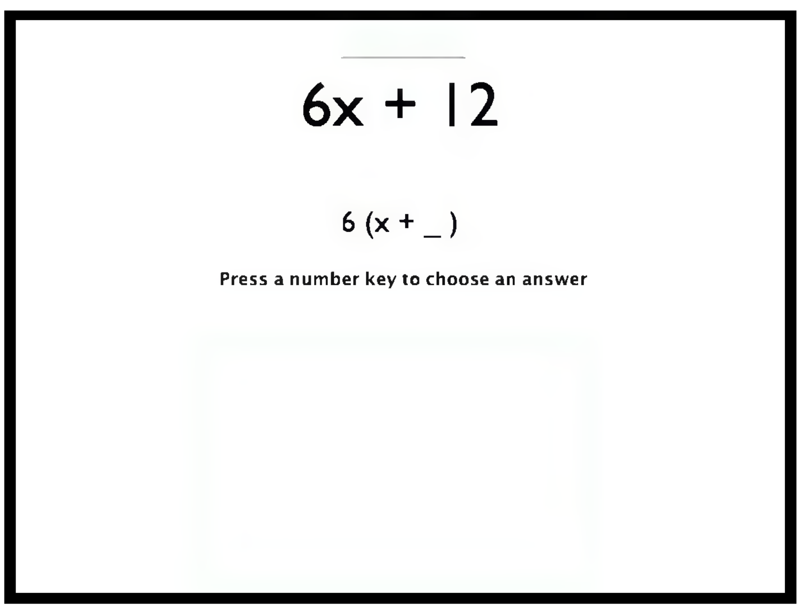


Figure 10. Example far transfer item slide: polynomial form Bx – C.

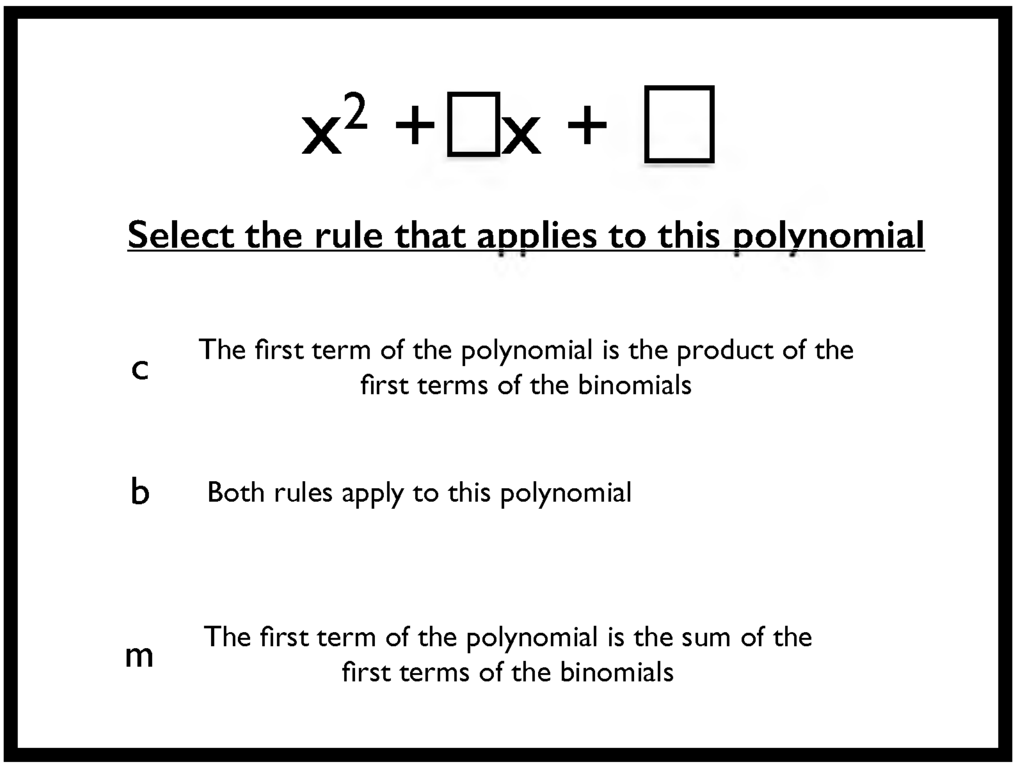


Figure 11. Example item from the declarative rules test.

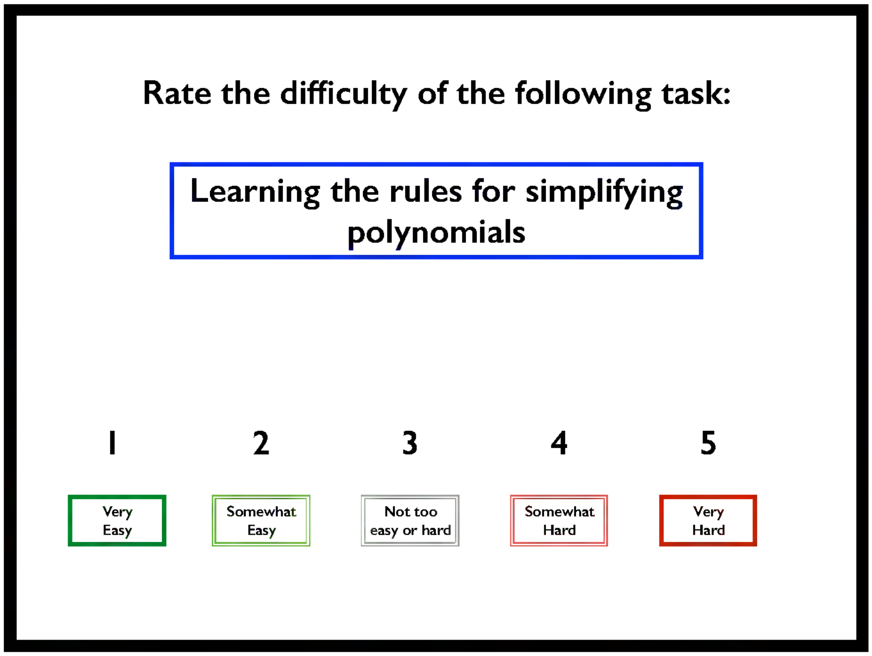


Figure 12. Example slide of Difficulty Questionnaire: rule learning difficulty.

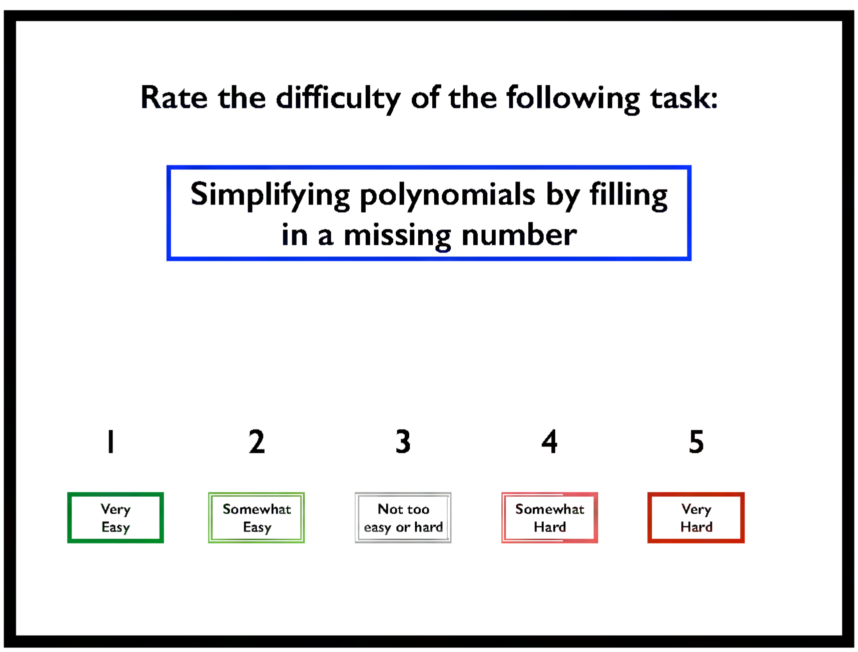


Figure 13. Example slide of Difficulty Questionnaire: problem solving difficulty.

CHAPTER 4

RESULTS

The following measures were obtained from the final sample of 171 participants over 3 days: proportion of errors and response time (RT) for the implicit learning trials and learning and transfer tests, proportion of errors for the declarative rules test, coding of participant responses to rule generation questions, overall time and slide count for the Declarative Learning Task, and difficulty ratings for the questionnaire. As noted in the previous chapter, subjects were randomly assigned to a learning condition within each classroom, but they were nested within teacher and school. In the analysis, the teacher was treated as a random factor but due to the potentially important procedural difference in the two schools, the school was treated as a fixed factor (i.e., the two schools represented a difference delay between learning events rather than a sample of schools receiving the same experimental conditions).

As noted previously, the sign patterns used in learning and near transfer, respectively, were counterbalanced in an attempt to avoid a confound. Analyses of the relevant study measures for effects of counterbalance group membership found no statistically significant differences. Table 4 shows the statistical test results for the main effect of the counterbalancing factor on each dependent measure. Given this outcome, this factor was dropped from the analyses.

Of the 188 participants who completed all tasks, 17 were eliminated due to unrealistically low RT on the three learning tests. The daily learning test was used to determine outliers because all participants took it on all days of the experiment, and I determined outlier thresholds using the absolute deviation around the median (Leys, Ley, Klein, Bernard, & Licata, 2013). Participants with average response times lower than the outlier threshold on any of the learning tests were excluded from data analysis. Because the latency data for the three learning tests were positively skewed, I used the natural log of each participant's response time in calculating the thresholds. The median absolute deviation, or MAD, is calculated using the median of the absolute distances of all scores from the sample median. This MAD is used similarly to a standard deviation to calculate an outlier threshold; I used an outlier threshold of 2.5 times the MAD away from the median, as recommended by Leys and colleagues (2013), which is considered moderately conservative. These values were then converted back from log RT to RT to obtain the appropriate outlier thresholds for each test.

The presentation of results and analyses in the remainder of this chapter are organized around the four research questions defined at the end of Chapter 2. The critical p value for all statistical tests was set at .05, and all eta squared values are partial eta squared. Cohen's d is also reported as an effect size estimate for appropriate analyses. For the repeated measures tests in which Mauchly's test indicated the assumption of sphericity had been violated, degrees of freedom were corrected using Greenhouse-Geisser estimates of sphericity.

Question One: Implicit Skill Learning

The hypothesis that participants could acquire polynomial factoring skills without exposure to the declarative rules was tested in two analyses. First, performance in the Implicit Learning Task was compared between the participants who first had 2 days of declarative instruction and those that had none. If the declarative-first group showed better performance on the implicit task, this would indicate that declarative knowledge prior to procedural practice was important. A lack of difference would suggest that a comparable degree of procedural skill could be acquired from the current method without prior declarative knowledge. Second, performance on the end-of-session Learning tests of the two groups was compared. Equivalent or superior performance by the implicit-first group would provide support for the implicit acquisition of procedural skill in the absence of declarative knowledge.

Implicit Learning Task

Errors and RT were analyzed with ANOVA with learning condition, teacher, and school as between groups factors (teacher as a random factor nested within school as a fixed factor) and nine blocks on Session 1 of the Implicit Learning Task as the within-subject factor. Figure 14 (first nine blocks only) shows mean percentage errors and RT by group. The pattern of decreasing RT over the first three blocks followed by an increase in the fourth block, with a similar increase from Blocks 6 to 7, is consistent with the changes of foil structure in this task. After each set of three blocks, or triad, the foils changed to be more similar to the correct answer. The increasing error rate within triads is consistent with the nature of the numerical components of the foils; the numbers in the foils were identical to the correct answer by the 3rd block of each triad. Chance error percentage on this task was 50%.

No differences between the overall means of declarative-first and implicit-first groups were found in RT (M = 2200.8 ms, SD = 593.9, and M= 2194.0 ms, SD = 539.5, respectively), $F(1, 169) < 1$,[1] or in error rate (M = 10.25%, SD = 8.00 and M = 10.98%, SD = 5.72, respectively), $F(1, 169) < 1$. There was no interaction between block and learning condition for error rate, $F(6.36, 1074.9) < 1$. There was a small but significant interaction for RT, $F(3.70, 629.93) = 9.35$, $p < .001$, $\eta^2 = .051$. As seen in Figure 14, this is likely due to the slightly slower initial performance of the implicit-first group on Block 1 and slightly faster performance over the final 6 blocks. Overall, the implicit-first and declarative-first learning conditions showed equivalent performance across their respective initial session of the Implicit Learning Task. It did not appear that declarative rule instruction was necessary for learning to occur in the implicit task. The implicit-first group performed the various problem types with equivalent speed and accuracy to those who had two days of declarative instruction beforehand.

Learning Tests: Sessions 1 & 2

Errors and RT were analyzed with ANOVA with learning condition, teacher and school as between groups factors (teacher as a random factor nested within school as a fixed factor) and test session as the within-subject factor. Two orthogonal contrasts, average and difference of the within-subject

[1] Mauchly's test indicated that the assumption of sphericity had been violated $x^2(35) = 672.87$, $p < .001$); therefore, degrees of freedom were corrected using Greenhouse-Geisser estimates of sphericity.

factor, were tested. Figures 15 and 16 show mean percentage error and RT, respectively, by session and group; chance performance for errors was 88.9% [2] on this task. There were main effects of learning condition for error rate, $F(1, 3) = 17.52$, $p = .024$, $\eta^2 = .852$, $d = 1.108$, and response time, $F(1, 3) = 80.72$, $p = .003$, $\eta^2 = .963$, $d = 1.86$, with the implicit-first learning condition outperforming the declarative-first learning condition on both measures. There was an interaction between session and learning condition for error rates, $F(1, 3) = 8.42$, $p = .048$, $\eta^2 = .692$, $d = .175$, which indicated that the error rates decreased more from Session 1 to Session 2 for the declarative-first group than for the implicit-first group.

There was also an interaction between condition and teacher, $F(3, 3) = 2.83$, $p = .04$, $\eta^2 = .050$, which appears to be driven by the participants from a single teacher; those participants had the highest error percentage in the declarative-first condition and the lowest error percentage in the implicit-first condition. There was also an interaction between day and teacher for error rates, $F(3, 3) = 18.30$, $p = .02$, $\eta^2 = .948$, which indicates that the error rate change from Session 1 to 2 differed by teacher, in this case, participants from two teachers with fewer students participating showed greater gains from

[2] This was a free response task, but because the correct answer is always a non-zero single digit, there are 9 possible responses.

Session 1 to 2. There was also a 3-way interaction for RT between day, condition and teacher, $F(3, 161) = 3.12$, $p = .028$, $\eta^2 = .055$. This interaction appears driven by an increase in average RT for one teacher's declarative-first learning from Session 1 to 2, while all other groups experienced RT decreases. There were no other interactions. Despite the presence of some teacher interaction effects, the Session 1 and 2 learning test data show overall lower error rates and response times for participants in the implicit-first learning condition than those in the declarative-first learning condition.

Reliability estimates

To estimate split-half reliability for the Session 1 and 2 learning tests, all participants' error rate scores were calculated separately for even and odd numbered trials. Using the Spearman-Brown adjustment, internal consistency reliability estimates for the learning tests were as follows, Session 1 $r_{xx'} = .906$, Session 2 $r_{xx'} = .904$.

Question Two: Rule Learning

The hypothesis that participants who received initial implicit training on polynomial factorization would more easily learn declarative rules for that skill compared to those who did not engage in implicit training was tested in two analyses. First, performance on two tests of declarative rule learning, a forced choice and a free response, was compared between those who had implicit training before learning the rules and those without implicit training beforehand. If the implicit learning first group performed better on the rule tests it would suggest that implicit training could benefit acquisition of declarative rules for that skill. Second, the amount of time spent, and number of slides viewed, on the declarative rule learning task was compared between those who had implicit training and those who did not. If the implicit learning first group spent less time or viewed fewer slides than those without the training, it would support the hypothesis that implicit training aids later declarative rule learning.

Declarative Rules Test

Error rates for the declarative rules pretest and posttest were analyzed using Univariate ANOVA analyses. Figure 17 shows pretest and posttest mean error rates by group; chance performance on this task was 66.6% errors. There were no learning condition group differences on the declarative rules pretest, $F(1, 3) < 1$, which suggests that the two groups had similar levels of explicit understanding of the rules prior to the Declarative Learning Task. There was, however, a significant difference between the two learning conditions on the posttest, $F(1, 3) = 12.58$, $p = .034$, $\eta^2 = .796$, $d = .339$. Participants in the implicit-first condition outscored those in the declarative-first condition by about 5 percentage points, (M = 61.53, SD = 14.43 and M = 66.28, SD 13.55, respectively). Both groups performed very poorly on this task, but those in the implicit-first group did score slightly better on the posttest, which provides little, if any, support to the hypothesis that initial implicit skill learning is beneficial to later rule learning.

Declarative Rule Generation

Participants generated rules for simplifying two types of polynomials before the declarative rules test. Participant responses were coded to reflect the total number of idea units each participant recorded, divided into total correct idea units, total incorrect idea units, and total unrelated idea units. Table 5 shows common participant examples that were coded as correct, incorrect, and unrelated idea units. Table 6 displays the means and standard deviations of the proportion of correct idea units and total idea units for the pretest and posttest for each group. The coded scores were analyzed by learning condition with ANOVA. In order to gauge the effectiveness of the scoring rubric, a second rater scored 20 participant responses. An intraclass correlation was calculated between the two raters, ICC = .894, $p < .001$. There were no significant differences in the proportion of correct idea units before or after the declarative learning, $F(1, 3) < 1$ and $F(1, 3) = 1.23$, $p = .284$, respectively. There were also no significant differences in the total idea units on the pretest, $F(1, 3) < 1$, but there was a significant difference on the posttest, $F(1, 3) = 8.25$, $p = .021$, $\eta^2 = .512$, $d = .417$, with participants in the declarative-first condition recording more idea units than those in the implicit-first condition. This evidence does not support the hypothesis that initial implicit training of problem patterns would facilitate the subsequent acquisition of declarative rules for problem solving.

Modified rule generation

Nineteen students participated in a modified version of the experiment in the implicit learning first group. Everything was identical for this group, except for the rule generation portion of the declarative rules test. Participants described how to teach another student how to solve two different types of polynomials. There were no correct idea units recorded by any participant prior to the Declarative Learning Task. This result suggests that implicit skill learning does not, in and of itself, result in a declarative understanding of the rules driving those skills.

Declarative Learning Time and Slide Count

Time spent on the Declarative Learning Task (recorded in minutes) and number of slides viewed were analyzed by learning condition with a one-way ANOVA. Participants in the implicit-first learning condition spent significantly less time (M = 11.44 minutes, SD = 1.21) than the participants in the declarative-first learning condition (M = 13.04 minutes, SD = 3.05), $F(1, 3) = 48.51$, $p = .004$, $\eta^2 = .935$, $d = .69$. As such, those in the declarative-first condition spent 12.8% more time on the Declarative Learning Task than those in the Implicit Learning Task, but scored slightly worse on the posttest. Participants in the implicit-first learning condition viewed fewer slides in the Declarative Learning Task than those in the declarative-first learning condition (M = 38.53, SD = 2.56 and M = 41.31, SD = 4.49, respectively), and the difference was significant, $F(1, 3) = 21.15$, $p = .018$, $\eta^2 = .872$, $d = .769$. Since participants could not skip slides, this difference indicates that those in the implicit-first learning condition went back to repeat slides fewer times than those in the declarative-first condition. The slide count data, along with the overall learning time, support the hypothesis that initial implicit training will benefit declarative rule learning.

Question Three: Perception of Difficulty

The hypothesis that participants who engaged in implicit skill training prior to declarative rule learning would perceive solving problems and learning declarative rules as less difficult than participants who learned the declarative rules first was tested in one analysis. Average responses on the difficulty perception questionnaire were compared between those who engaged in implicit learning first and those who learned declarative rules first. If the implicit-first group indicated they perceived the rule learning and problem solving as less difficult than those who learned declarative rules first, it would support the hypothesis that initial implicit training positively impacts students' perceptions of difficulty.

Difficulty Questionnaire

Participants completed 2 questions, one involving the difficulty of solving polynomials, and one involving the difficulty of learning the rules for solving the polynomials. Results were analyzed with ANOVA. Table 7 displays means and standard deviations for both questions. There were no significant differences between the learning condition groups for either the problem solving difficulty or the rule learning difficulty, $F(1, 3) < 1$ and $F(1, 3) = 1.63$, respectively, nor were there any significant teacher effects or interactions. While there were no significant differences, the results were trending towards participants in the implicit-first learning condition perceiving both the problem solving and rule learning as more difficult than those in the declarative-first condition.

Question Four: Problem Solving and Transfer

The hypothesis that initial implicit skill training before declarative rule learning will lead to better problem solving and transfer performance was tested in one analysis. Performance on a final day Learning Test, Near Transfer Test and Far Transfer Test was compared between participants who had implicit learning first and those who had declarative learning first. If the implicit learning first group showed better performance on these tests it would support the hypothesis that initial implicit skill learning leads to better acquisition and transfer of procedural skills as compared with initial declarative rule learning.

Day 3 Tests: Learning and Transfer

Errors and RT were analyzed with ANOVA with learning condition, teacher and school as between groups factors (teacher as a random factor nested within school as a fixed factor) and test type (learning, near transfer & far transfer) as the within-subject factor. Three orthogonal contrasts were tested for the within-subject factor: average of the 3 tests, the contrast of the learning test and the combined transfer tests, and the contrast of the near and far transfer tests. Figures 15 and 16 also show error and RT means for the Session 3 learning, near and far transfer tests labeled as 3A, 3B and 3C, respectively. Chance performance on these tasks was 88.9% errors.

There was a main effect of learning condition on RT across the three measures, $F(1, 3) = 26.36$, $p = .006$, $\eta^2 = .862$, $d = .285$, and the implicit-first group had faster response times than those in the declarative-first condition, but there was no main effect for errors, $F(1, 3) < 1$. There was also a significant main effect of teacher on RT, $F(3, 3) = 32.06$, $p = .009$, $\eta^2 = .970$, which indicates that participants with different teachers had different average response times. There were no significant effects for the learning test versus transfer tests comparison for RT or errors, $F(1, 3) < 1$, and $F(1, 3) = 1.21$, $p = .366$, respectively, or for the near transfer test versus far transfer test comparison for RT or errors, $F(1, 3) = 1.64$, $p = .288$ and $F(1, 3) < 1$, respectively. Overall, the results of the final learning and transfer tests add little support for the hypothesis that initial implicit training results in better final problem solving and transfer as compared to declarative learning first. The only difference between the two groups was that participants in the implicit-first condition were somewhat faster than those in the declarative-first condition on the final day tests.

Reliability estimates

To estimate split-half reliability for the Session 3 learning and transfer tests, all participants' error rate scores were calculated separately for even and odd numbered trials. Using the Spearman-Brown adjustment, internal consistency reliability estimates for the learning and transfer tests were as follows, Session 3 learning $r_{xx'} = .937$, near transfer $r_{xx'} = .955$, and far transfer $r_{xx'} = .601$. The reliability estimates for the learning and near transfer indicate consistent items, but the estimate for the far transfer test was lower despite similar levels of accuracy.

Table 4. Results of tests of statistical significance of counterbalance group membership.

Measure	F Value
Declarative Rule Tests: Error Percentage	$F(2, 168) = 2.179, p=.116$
Session 1 & 2 Learning Tests: RT	$F(2, 168) = 2.243, p=.109$
Session 1 & 2 Learning Tests: Error Percentage	$F(2, 168) = 1.221, p=.297$
Session 3 Tests: RT	$F < 1$
Session 3 Tests: Error Percentage	$F < 1$
Declarative Learning: Overall Time	$F < 1$
Declarative Learning: Slide Count	$F < 1$
Difficulty Questions	$F < 1$

Table 5. Examples of participant responses to declarative rule generation as coded into correct, incorrect, and unrelated idea units.

Correct Idea Units	**Incorrect Idea Units**	**Unrelated Idea Units**
distributive property	To multiply use parentheses?	3x
first write parentheses	Multiply the number by the exponent	2x + 0
find the greatest common factor	You add the monomials together	2 + 2
	Add the variables	2x
		I don't know

Table 6. Means and standard deviations for coded participant responses to declarative rule generation.

Coded Units	**Condition**	**Mean Pre**	**SD Pre**	**Mean Post**	**SD Post**
Proportion Correct	Declarative First	.00	.00	.11	.17
	Implicit First	.01	.07	.11	.19
Total Idea Units	Declarative First	1.75	1.33	2.48	1.56
	Implicit First	1.69	.92	1.92	1.08

Table 7. Means and standard deviations for Difficulty Questionnaire.

Question	Condition	Mean	SD
Problem Solving Difficulty	Declarative First	2.93	1.21
	Implicit First	3.20	1.40
Rule Learning Difficulty	Declarative First	3.11	1.23
	Implicit First	3.52	1.23

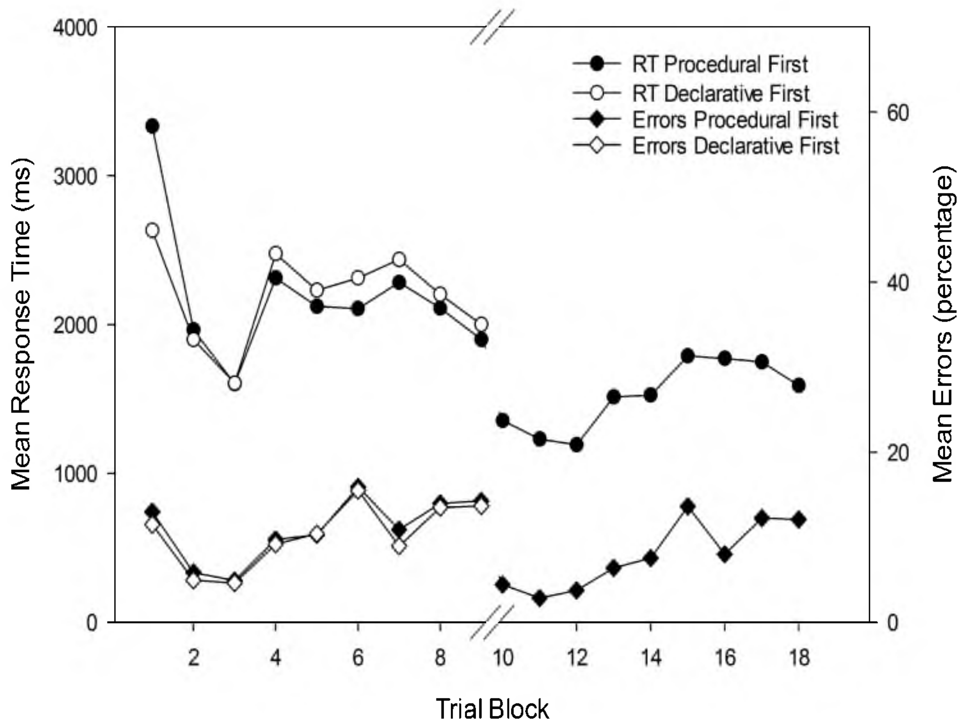


Figure 14. Mean RT and percent error for Implicit Learning Task trials by group. Note: Blocks 1–9 are included for both groups with the implicit-first group receiving them on Session 1 and the declarative-first group receiving them on Session 3. Blocks 10–18 represent the nine blocks on the second training session for the implicit-first learning group.

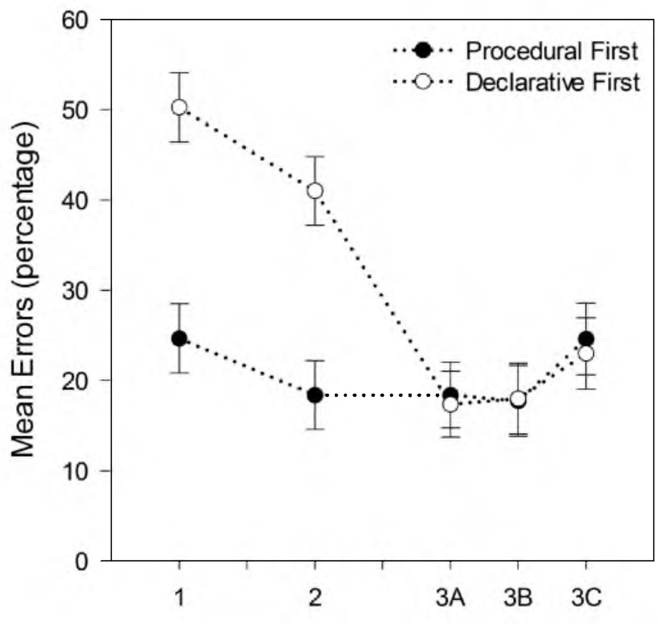
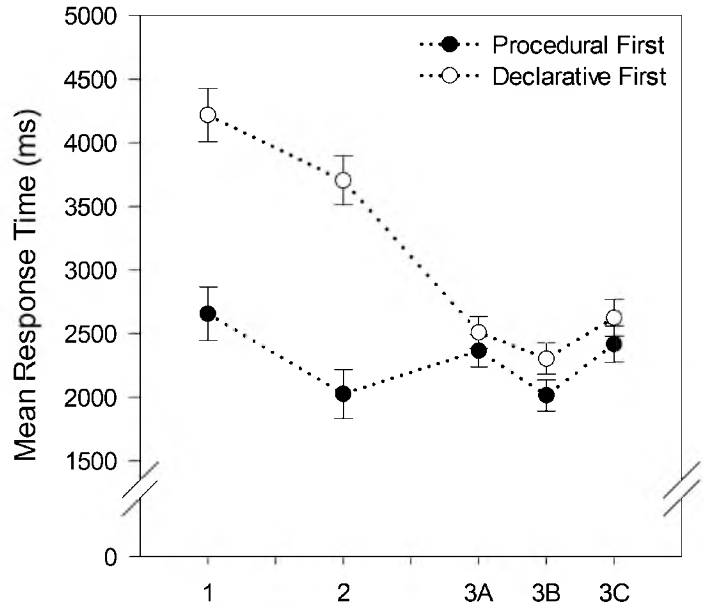


Figure 15. Mean percent error for learning tests and transfer tests by group. Note: 1, 2, and 3A are learning tests from Sessions 1, 2, and 3, respectively. 3B and 3C are the near and far transfer tests, respectively. The error bars represent 95% confidence intervals.



Test Session

Figure 16. Mean RT for learning tests and transfer tests by group. Note: 1, 2, and 3A are learning tests from Sessions 1, 2, and 3, respectively. 3B and 3C are the near and far transfer tests, respectively. The error bars represent 95% confidence intervals.

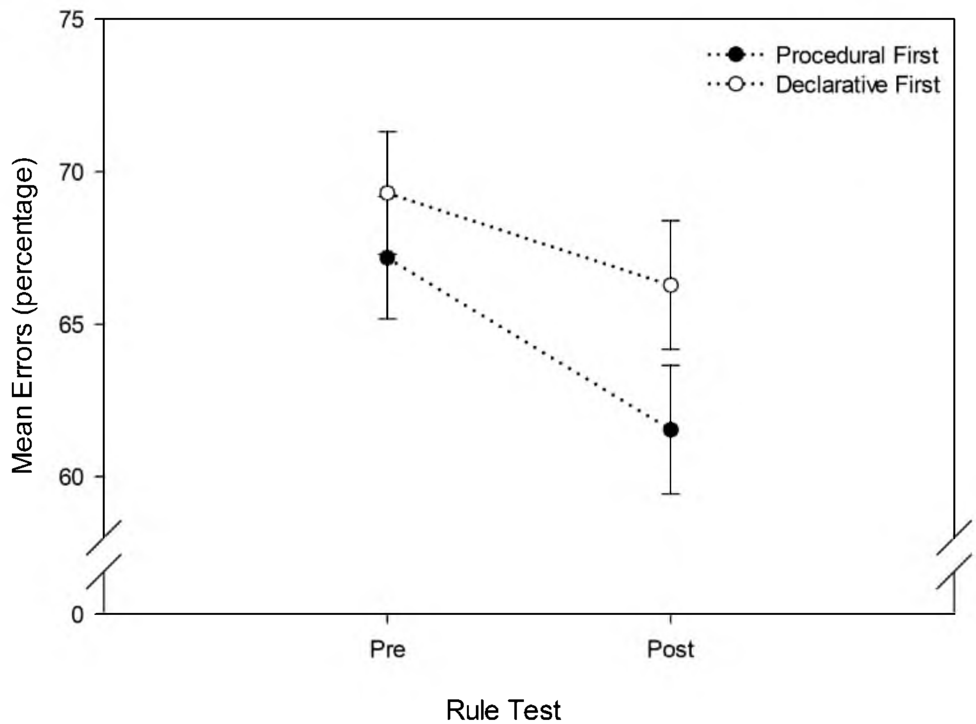


Figure 17. Mean percent error for the declarative rules test given before and after the Declarative Learning Task. Note: The error bars represent 95% confidence intervals.

CHAPTER 5

DISCUSSION

The present experiment investigated the possibility of learning algebraic problem solving skills through initial implicit training, absent of any declarative training. In the event this implicit training was found to promote skill learning, the experiment investigated the impact of that training on rule learning, difficulty perception and final problem solving and transfer. The implicit training consisted of practice recognizing patterns in both problem structure and numeric relationships that are associated with correct polynomial factoring solutions. The practice was designed such that participants' responses to the pattern matching exercises were relatively error free. Varied skill and rule tests and other measures were employed to investigate what impact, if any, this implicit skill training had on the learning of this skill when presented prior to declarative instruction. The results of analyzing error rates and response times support the hypothesis that some implicit knowledge of problem patterns in polynomial factoring can be learned without declarative instruction. Evidence for this conclusion and evidence regarding the impact of this learning will be discussed with respect to the four original research questions.

Question One: Can Algebraic Problem Solving Skills Be Acquired Implicitly?

The current evidence supports the feasibility of implicitly acquiring some polynomial factoring skill in the absence of explicit instruction in the declarative rules. Participants in the implicit-first learning condition performed as well on the Implicit Learning Task as those in the declarative-first learning condition who had trained on the rules for solving those problems prior to the learning task. In this case, there was no benefit to learning the rules prior to implicit skill training. This could be attributed in part to the design of the Implicit Learning Task that promoted relatively error-free performance regardless of background knowledge.

In contrast to equivalent performance by the two groups during implicit training, participants in the implicit-first learning condition outperformed the participants in the other group on the first two daily learning tests, which occurred prior to participants' exposure to the learning task they had previously not encountered. Implicit-first participants committed fewer errors and responded more quickly to the items than those in the declarative-first condition. The end-of-session learning tests required participants to generate a numeric response to a missing element in a factored polynomial, something that neither group had to do during training. The declarative-first group had been exposed to examples and rule explanations for solving such problems, and the implicit-first group had practiced selecting pattern solutions to such problems in a forced choice format. It seems likely that this finding is at least partially due to the robustness of implicit learning mechanisms to high working memory demands, such as those required to learn to solve a complex skill as this. In all, the implicit-first learning group performed equally or better than the declarative learning first group on all tasks prior to their declarative instruction, suggesting that the partial acquisition of problem solving skill implicitly without declarative instruction is possible and has some potential benefit.

Question Two: Does Initial Implicit Learning of Algebraic Problem Solving Skills Increase the Ease of Learning the Associated Declarative Rules?

Analysis of data from the declarative rules test and declarative rule learning task provided little if any support for the hypothesis that initial implicit learning of algebraic problem solving skills would improve declarative rule learning. Participants in the implicit-first learning condition did not differ from those in the declarative-first condition on the declarative pretest, even though they performed it after 2 days of implicit training that exposed them to problem and solution patterns. This finding suggests that the knowledge gained during the implicit training, knowledge that was demonstrated in better end-of-session learning test performance compared to that of declarative-first participants, was not explicit understanding of polynomial factoring. It appeared to be implicit understanding of problem structure and solution patterns. On the declarative knowledge posttest, mean performance of implicit-first participants was marginally better (fewer errors) than those in the declarative-first learning condition. This is consistent with the prediction that initial implicit understanding of problem structure can facilitate the acquisition of declarative rules for problem solutions. However, the difference was small, and both groups' response accuracy was close to chance.

Analysis of the number of Declarative Learning Tasks slides viewed and overall time on that task provided somewhat stronger support for the hypothesis that implicit training would facilitate subsequent declarative learning. Participants in the implicit-first learning condition spent significantly less time and viewed fewer slides than those in the declarative-first learning condition. Because the implicit-first group experienced their single session of declarative instruction after two days of participating in the experiment, and the declarative-first group experienced their first session of declarative instruction on Session 1, the time and slide differences could reflect motivational differences associated with familiarity versus novelty of the context. However, this alternative explanation is inconsistent with the RT data for the first session of implicit training. Declarative-first participants were not faster than implicit-first participants, despite the fact that they performed this task on Session 3 compared to Session 1. It therefore seems more plausible that the reduced declarative learning time by the implicit-first group reflected facilitation from prior implicit knowledge of problem structure. Nevertheless, the reduced learning time and number of slides viewed by the implicit-first group cannot be viewed as compelling evidence for facilitation from their implicit training in light of their near chance accuracy on the declarative knowledge posttest.

Analysis of the coded data from the responses participants generated in response to being asked how to simplify the polynomials yielded less supportive results, as there were no differences between the two groups. Both groups generated very low proportions of correct ideas, which may speak to the difficulty of learning complex rules and the quality of the rule knowledge participants were able to internalize. In any event, initial implicit skill learning did not provide any benefit to the generation of problem solving rules.

Question Three: Does Initial Implicit Learning of Algebraic Problem Solving Skills Impact Perceptions of Difficulty in Problem Solving or Skill Learning?

Student perceptions of difficulty in learning rules and solving problems were reported on a Likert-like scale and, while indicating no significant differences, showed trends toward participants in the implicit-first learning condition finding both the rule learning and problem solving more difficult than those in the declarative-first condition. This evidence is contrary to what was hypothesized, as well as to some aspects of the performance data. The difficulty rating scale was given at the end of Session 3 of the experiment, so it is possible that these difficulty ratings more accurately reflected the participants' perceptions of difficulty on that session's tasks rather than perceived difficulty of the entire learning experience. If that is the case, the participants' reports could simply indicate that the declarative rule-learning task is perceived as more difficult than the implicit rule-learning task. This alternative explanation for the ratings would be consistent with the time participants spent on the Declarative Learning Task and how many slides they viewed, as the participants in the implicit-first learning condition spent less time and viewed fewer slides than those in the declarative-first condition. Given that these participants performed slightly better on the rule tests, the decreased time on the learning task and fewer times repeating slides seems to support the idea that declarative learning was no more difficult and perhaps easier for them. The implicit-first learning participants also responded more quickly on the Session 1 and 2 learning tests and the near transfer test, and trended that way on the Session 3 learning and far transfer tests. Despite the performance data suggesting ease of learning in the implicit-first condition, the difficulty rating data cannot be discounted unless future research provides contrary evidence when ratings are obtained at each session rather than only on the final session.

Question Four: Does Initial Implicit Learning of Algebraic Problem Solving Skills Impact Final Problem Solving and Transfer Skill?

The analyses of the learning and transfer tests at the end of Session 3 of the experiment indicated no differences on error rates due to learning condition, which is inconsistent with the hypothesis that initial implicit training would lead to better problem solving and transfer the skills. There was a significant condition effect on RT in Session 3 tests, with the implicit-first group responding more quickly than the declarative-first group, which supports the hypothesis, as speed is one element of effective problem solving skill. Despite this piece of evidence, the larger body of evidence of Session 3 tests supports the notion that initial implicit algebraic skill learning has no effect on final problem solving and transfer outcomes.

Related to the lack of clear performance differences on Session 3, two observations are worth noting. First, the end-of-session learning test performance by the declarative-first group was relatively slow and inaccurate on their first 2 days of instruction. However, a single session of implicit training resulted in learning test performance equivalent to the implicit-first group. This presumably attests to the effectiveness of the implicit, error-free exposure to problem structure in the current test of polynomial factoring. Second, in both groups there was little performance decline on the transfer tests compared to the final learning test (refer to performance by both groups on Tests 3B and 3C relative to 3A in Figures 12 and 13). This is despite the fact that the transfer tests presented new sign patterns in the case of near transfer and entirely new types of polynomials in the case of far transfer. Given the relatively poor performance on the learning tests by the declarative-first group prior to implicit training, this transfer performance is likely dependent to a large extent on the implicit exposure to problem and solution patterns. Although procedural knowledge is often described as hyperspecific and resistant to transfer, this evidence suggests substantial transfer of implicit knowledge.

Given the differences between the learning items and the near and far transfer items, it seems likely that transfer of a generalized rule was required, as opposed to a superficial transfer. Figure 18 shows a polynomial and factorization from the learning, near transfer, and far transfer sections. While the near transfer items are similar to the learning items, the sign patterns of the solutions are completely different. Thus, a generalized rule must be applied to solve these problems, as patterns implicitly learned during training would not specifically apply to these problems. The far transfer items required further rule generalization; participants needed to realize that factor pairs that sum to zero would produce the absence of the middle term of the polynomial in order to simplify the polynomial.

Implications for Skill Acquisition Literature

Declarative Knowledge and Implicit Learning

There are enduring questions in the implicit learning and skill acquisition literatures pertaining to the role of declarative knowledge in implicit skill learning. One debate is whether declarative rules for an implicitly learned skill can become spontaneously acquired through the implicit learning process. Evidence from Willingham et al. (1999) using the serial response task indicated the possibility of such spontaneous rule learning.

Evidence from the present study indicated no such effect in the more complex domain of polynomial factoring, as virtually no correct idea units were recorded in a declarative rule generation task for participants in the implicit learning condition, and on average these participants scored at chance on the multiple choice declarative rules tests. Chance performance on the multiple choice measure suggests that even a more sensitive recognition memory measure cannot produce evidence for an explicit understanding of polynomial factoring rules in this task, even though participants were clearly able to use an implicit understanding to solve problems quickly and accurately.

Another debate is whether declarative rule knowledge is necessary for skill proceduralization. Anderson (1983, 1992) and colleagues (1994, 1997, 2004) have consistently posited that declarative knowledge must precede proceduralization of skills. Research by others (Nissen & Bullemer, 1987; Willingham et. al, 1989; Nissen, 1992; Reber & Squire, 1998) suggests that procedural skills can be learned in absence of declarative knowledge. The present study provides evidence supporting the latter research, as participants learned to solve problems with only implicit training and demonstrated little if any understanding of the declarative following this training.

Implicit learning domains

Much of the existing research on implicit learning has investigated motor skill learning (Maxwell et al., 2001; Mary Jo Nissen, 1992; Willingham & Goedert-Eschmann, 1999) or cognitive skills in the linguistic domain (Hartman et al., 1989; Kessels & de Haan, 2003; A. S. Reber, 1967; Warmington et al., 2013; Wessel et al., 2012). Previous research on cognitive implicit learning has typically been focused on learning novel words or grammars. The literature in artificial grammar learning and literature on implicit learning of new words or word associations has provided evidence for the feasibility of language related patterns and knowledge to be learned in the absence of declarative rules or training. However, it could be implied that this body of evidence represents a unique phenomenon that is limited to the domain of language learning.

The present experiment provides evidence to the contrary of that assertion. Participants in the study were, with no declarative rules or training, able to implicitly learn the patterns involved in complex polynomial simplification such that they were able to accurately simplify novel polynomials quite quickly. This, in and of itself, provides evidence that implicit cognitive learning may not be a phenomenon limited to natural language processing abilities. Providing further evidence to that claim is the lack of declarative understanding shown by participants in the implicit learning condition prior to declarative rule learning. Given that participants were unable to identify rules for simplifying polynomials, provide an explanation for how to simplify polynomials, or explain how to teach someone how to simplify polynomials, it suggests that participants had little or no explicit understanding for simplifying the polynomials they were proficient at solving. If participants had no explicit understanding of how to solve the polynomials, yet were able to solve them effectively, it suggests implicit learning of mathematical problem solving skills, a domain outside of the linguistic domain heavily studied in implicit learning literature.

Complex skill acquisition

Anderson's work on skill acquisition has explored the processes involved in the acquisition of more complex cognitive skills such as computer programming (Anderson, Conrad, & Corbett, 1989) and other complex skills (Anderson & Fincham, 1994; Anderson, Fincham, & Douglass, 1997). This research has investigated the use of examples in skill acquisition and concluded that an analogy mechanism is likely used in addition to direct recall of problems (Anderson & Fincham, 1994; Anderson, Fincham, & Douglass, 1997). Anderson and colleagues determined that examples were encoded declaratively and used to aid early problem solving when applying declarative rules. Schwartz and

Bransford's (1998) study investigated the principles of problem exposure prior to instruction in a classroom setting and also found it was beneficial to the quality of learning. The findings of the present study are consistent with these studies on skill acquisition, particularly utilizing problem and example exposure prior to other learning. The present study differs, however, in that it was designed to explore the plausibility of implicit learning from multiple problem exposures and its impact on later declarative learning and skill application.

Implications for Mathematics Literature

There are many studies in the mathematics learning literature seeking means of improving mathematics skill acquisition. Many of these methods even seek to mitigate the limitations of working memory on complex mathematics skill learning (Ashcraft & Krause, 2007; Atkinson et al., 2000; Atkinson & Renkl, 2007; DeCaro & Rittle-Johnson, 2012; Renkl et al., 2004; Renkl & Atkinson, 2007). Despite this similarity with the goals of the present study, the referenced studies relied on explicit, declarative processes for initial learning in the domain. The present study succeeded in separating implicit and explicit processes and provided evidence for the possibility of learning mathematics problem solving skills implicitly. In addition, there was limited support for the hypothesis that initial implicit skill learning could decrease demands on working memory and increase skill and rule learning in the math domain. Given the findings from the present study that indicate the plausibility of implicitly learning complex mathematical skills, but marginal and conflicted results in many other areas, a reasonable course of action would be to investigate how implicit skill learning in math could be utilized to improve learning. Research in mathematics problem exploration (DeCaro & Rittle-Johnson, 2012) and worked examples (Renkl et al., 2004) have shown benefits of exposure to domain problems prior to math instruction; these methodologies may show increased learning benefits from the addition of some form of initial implicit training prior to the other forms of domain exposure, and the combination may provide more information about the nature, implicit or declarative, of the benefits derived from worked examples or problem exploration prior to instruction.

Limitations

Although the present investigation yielded findings that extend existing evidence in the implicit learning literature, it had several important limitations. First, no claims can be made for the comparative effectiveness of the implicit and declarative training tasks. Both were designed to promote distinct types of knowledge and memory representations: procedural and declarative. Undoubtedly, neither was optimal in achieving this. The implicit training task was modeled on the errorless learning paradigms used to train individuals with declarative learning deficits, and an attempt was made to omit all declarative expressions of the algebraic rules being taught. The declarative instruction was designed to reflect the content of a popular algebra textbook, with some use of multimedia enhancements to facilitate comprehension of the explanations. Despite efforts to make each instructional condition as effective as possible given the constraints imposed by their intended promotion of either declarative or procedural knowledge exclusively, there is no way to judge their relative effectiveness. The end-of-session learning tests indicated that the implicit training produced better outcomes. This could reflect the fact that declarative learning of procedures for polynomial simplification place unmanageable demands on working memory regardless of the instructional design. Or, this outcome could reflect instructional methods in the Declarative Learning Task that were farther from optimal than those in the Implicit Learning Task.

Related to this, another limitation of this experiment is the ambiguity of whether the end-of-session learning tests and the transfer tests had greater overall similarity with, and therefore favored, the Implicit Learning Task. If this were the case, then the observed differences may be more related to a task familiarity than reduction of working memory demands during learning. However, the tests were designed to differ from the Implicit Learning Task in several ways. First, the learning tests required participants to complete a partial factorization; this is quite different from the Implicit Learning Task, in which participants selected the correct answer from two complete factorizations. Selecting which of two possible factorizations is correct is likely a different process than completing a factorization; it is not dissimilar from the difference between recognizing and recalling correct answers. Second, the learning test was not a forced choice, as participants were required to fill-in-the-blank with a number, while the Implicit Learning Task only required participants to choose from among two answers. Presumably, these were large enough differences to eliminate or minimize any effects of task familiarity on the outcome of the learning tests. In addition, the Declarative Learning Task presented participants with the steps to factoring a polynomial using example problems; as such, participants would have seen how to fill in missing values in a solution. Ultimately, however, it is still possible that the solving problems in the Implicit Learning Task provided more familiarity with the learning tests, and that is what drove the differences in Session 1 and 2 learning test performance.

Another related issue pertains to the comparability of the declarative instruction condition and typical classroom instruction. The researcher's observations about common classroom instruction and current algebra textbooks suggest an emphasis on declarative understanding prior to problem solving practice, and this was a primary motivation for the research questions pursued here. Although the current experiment's declarative instruction closely followed a popular textbook's content with attempts to present this material in an effective manner, this learning task does not represent how instruction occurs in a typical mathematics classroom. Textbook reading and verbal classroom instruction would be coupled with skill practice, and teacher clarification generally would be available during the problem solving practice. In this study, the declarative rule learning task and the structure of the experiment were designed to compare the theoretical differences between implicit skill learning before declarative rule learning versus declarative rule learning before implicit skill practice. Therefore, no comparisons can be drawn between declarative-first instruction in this experiment and typical classroom instruction that might place an initial focus on declarative understanding prior to problem solving practice. However, the Declarative Learning Task was the same for both the implicit-first learning group and the declarative-first learning group, and it was a manipulation of order of learning tasks that was being tested. Given that, the results can be interpreted from a theoretical perspective, even though the conditions are not comparable to the classroom setting in which mathematics is not solely declarative or implicit.

The lack of any measures of delayed retention in the study is also worth noting. There are a number of studies that utilize delayed retention tests as a means of measuring instructional manipulations, presumably because retention is a different construct than immediate skill acquisition. As such, it is possible that we could have gained a better understanding of the effects of initial implicit training on the quality of learning had there been a measure of delayed retention.

In addition, it is possible that there was increased difficulty during the Implicit Learning Task related to the strategy utilized in ordering the foils (see Figure 5). For each group of 3 blocks, 1–3, 4–6, and 7–9, the foil types were selected to be progressively more similar to the correct response; it is possible that an alternate order of foils types would have better represented a pattern of increasing similarity between correct and incorrect answer choices.

Another limitation of the present study relates to the Difficulty Questionnaire. Because this questionnaire was presented to participants at the end of the final day, it is unclear what was being rated. During the final session, participants completed the task that was different from what they performed during the previous two sessions. Consequently, it is possible they were rating the difficulty of the final day's task rather than the entire learning experience. In hindsight, difficulty perception data should have been collected at the end of each day, or perhaps even after the completion of each task. This would provide clearer information about participants' perception of difficulty of each learning condition.

A final limitation of the task is likely a limitation of many experimental studies. The present study only investigated the implicit learning paradigm to a single type of mathematical problem solving skill, and while many math problem-solving skills may be similar in nature, they are not the same. As such, it is not possible to generalize to all types of mathematical problem solving skills without further study.

Future Research

The present study was an initial attempt at understanding the impact of implicit learning in the mathematics domain. Based on the investigation itself, and the limitations of the present study, there are several areas in which future research should focus.

As previously noted, the Difficulty Questionnaire could be redesigned and more strategically located within the experimental tasks in order to more effectively understand the participants' perceptions of difficulty throughout the experiment. This is important given the research on mathematics anxiety (Ashcraft & Kirk, 2001; Ashcraft & Krause, 2007), in which anxiety can play a role in diminishing already taxed working memory resources during new math learning.

It would also be reasonable to explore the potential implications of initial implicit mathematics skill learning in real mathematics classrooms. There would be obvious limitations to such a study, such as the lack of random assignment to learning condition within a classroom and likely problems of low statistical power, but there would be clear benefits to the ecological validity of any findings. If, before a unit of instruction on polynomial simplification, experimental classes were given implicit training similar to that in the present study, and control classes were not, it would be possible to explore the impact of this implicit training on real classroom math teaching and learning. Since the participants would be receiving legitimate classroom instruction, it would allow conclusions to be drawn about any effect initial implicit training has on real world student learning.

There could also be value in exploring the possible impact of initial implicit skill training on the use of worked examples in mathematics learning. The worked examples approach to skill learning has been successful in reducing working memory load and increasing student performance. Could adding elements of implicit training prior to worked example exposure further reduce working memory demands and further increase learning benefits? It seems possible that a student possessing an implicit understanding of problem and solution patterns might be able to make greater use of worked examples of those problem types. This may even provide some bridging to self-generation of declarative rules that has been described in the implicit learning research (Willingham, Nissen, & Bullemer, 1989). Ultimately, it seems quite reasonable that implicit learning of procedural knowledge could dovetail with research on worked examples in mathematics.

Finally, future pursuit of these questions would need to apply the principle of initial implicit training to other types of math skills and age levels. Given the gradual development of working memory ability from childhood to early adulthood, it is possible that the impact of initial implicit learning would be greater at lower grade levels. Future research could test these issues at a lower age level with more basic problem solving skills. Furthermore, as the current study only utilized a single, albeit somewhat complex math skill, it will be important to verify that this training can be effective for a variety of types of math skills. Further research along this line could also move into more complex math skills, such as solving mathematical word problems.

Learning	Near Transfer	Far Transfer
$x^2 + 10x + 16$	$x^2 + 6x - 16$	$x^2 - 16$
$(x \quad)(x \quad)$	$(x \quad)(x \quad)$	$(x \quad)(x \quad)$
Factors: $8 \times 2 = 16$ $8 + 2 = 10$	Factors: $8 \times -2 = -16$ $8 + -2 = 6$	Factors: $4 \times -4 = -16$ $4 + -4 = 0$
Solution	*Solution*	*Solution*
$(x + 8)(x + 2)$	$(x + 8)(x + -2)$	$(x + 4)(x + -4)$

Figure 18. Partially worked example problems for learning, near transfer and far transfer.

REFERENCES

Ackerman, P. L. (1988). Determinants of individual differences during skill acquisition: Cognitive abilities and information processing. *Journal of Experimental Psychology: General, 117*(3), 288–318. doi: 10.1037/0096-3445.117.3.288

Ackerman, P. L. (1992). Predicting individual differences in complex skill acquisition: Dynamics of ability determinants. *Journal of Applied Psychology, 77*(5), 598–614. doi: 10.1037/0021-9010.77.5.598

Anderson, J. R. (1982). Acquisition of cognitive skill. *Psychological Review, 89*, 369–406.

Anderson, J. R. (1993). *Rules of the mind*. Hillsdale, NJ: L. Erlbaum Associates.

Anderson, J. R., Bothell, D., Byrne, M. D., Douglass, S., Lebiere, C., & Qin, Y. (2004). An integrated theory of the mind. *Psychological Review, 111*(4), 1036–1060. doi: 10.1037/0033-295X.111.4.1036

Anderson, J. R., Conrad, F. G., & Corbett, A. T. (1989). Skill acquisition and the LISP tutor. *Cognitive Science, 13*(4), 467–505.

Anderson, J. R., & Fincham, J. M. (1994). Acquisition of procedural skills from examples. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20*(6), 1322–1340.

Anderson, J. R., Fincham, J. M., & Douglass, S. (1997). The role of examples and rules in the acquisition of a cognitive skill. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 23*(4), 932–945.

Anderson, J. R., Reder, L. M., & Lebiere, C. (1996). Working memory: Activation limitations on retrieval. *Cognitive Psychology, 30*(3), 221–256. doi: 10.1006/cogp.1996.0007

Anderson, N. D., & Craik, F. I. M. (2006). The mnemonic mechanisms of errorless learning. *Neuropsychologia, 44*(14), 2806–2813. doi: http://dx.doi.org/10.1016/j.neuropsychologia.2006.05.026

Ashcraft, M. H., & Kirk, E. P. (2001). The relationships among working memory, math anxiety, and performance. *Journal of Experimental Psychology: General, 130*(2), 224–237. doi: 10.1037/0096-3445.130.2.224

Ashcraft, M. H., & Krause, J. A. (2007). Working memory, math performance, and math anxiety. *Psychonomic Bulletin & Review, 14*(2), 243–248.

Atkinson, R. K., Derry, S. J., Renkl, A. & Wortham, D. (2000). Learning from examples: instructional principles from the worked examples research. *Review of Educational Research, 70*(2), 181–214. doi: 10.2307/1170661

Atkinson, R. K., & Renkl, A. (2007). Interactive example-based learning environments: Using interactive elements to encourage effective processing of worked examples. *Educational Psychology Review, 19*(3), 375–386. doi: 10.1007/s10648-007-9055-2

Baddeley, A., & Wilson, B. A. (1994). When implicit learning fails: Amnesia and the problem of error elimination. *Neuropsychologia, 32*(1), 53–68. doi: http://dx.doi.org/10.1016/0028-3932(94)90068-X

Bell, M. (1998). *Everyday mathematics*. Chicago: Everyday Learning Corporation.

Bumby, D. R., Klutch, R. J., Collins D., & Egbers, E. (1995). *Integrated mathematics* (Teacher's Annotated ed.). New York: Glencoe/McGraw-Hill.

Charles, R. I., Branch-Boyd, J. C., Illingworth, M., Mills, D., & Reeves, A. (2004a). *Prentice Hall mathematics* (Teacher's ed., Course 3). Needham, MA: Prentice Hall.

Charles, R. I., Branch-Boyd, J. C., Illingworth, M., Mills, D., & Reeves, A. (2004b). *Prentice Hall mathematics: Study guide & practice workbook*. Needham, MA: Prentice Hall.

Cohen, N. J., & Squire, L. R. (1980). Preserved learning and retention of pattern-analyzing skill in amnesia: Dissociation of knowing how and knowing that.

Science, 210(4466), 207–210. doi: 10.1126/science.7414331

Corr, P. J. (2003). Personality and dual-task processing: Disruption of procedural learning by declarative processing. *Personality and Individual Differences, 34*(7), 1245–1269. doi: 10.1016/S0191-8869(02)00112-5

DeCaro, M. S., & Rittle-Johnson, B. (2012). Exploring mathematics problems prepares children to learn from instruction. *Journal of Experimental Child Psychology, 113*(4), 552–568. doi: 10.1016/j.jecp.2012.06.009

Dienes, Z., Broadbent, D., & Berry, D. C. (1991). Implicit and explicit knowledge bases in artificial grammar learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 17*(5), 875–887. doi: 10.1037/0278-7393.17.5.875

Evans, J. J., Levine, B., & Bateman, A. (2004). Research digest: Errorless learning. *Neuropsychological Rehabilitation, 14*(4), 467–476.

Fahle, M., & Daum, I. (2002). Perceptual learning in amnesia. *Neuropsychologia, 40*(8), 1167–1172. doi: 10.1016/S0028-3932(01)00231-7

Hartman, M., Knopman, D. S., & Nissen, M. J. (1989). Implicit learning of new verbal associations. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15*(6), 1070–1082. doi: 10.1037/0278-7393.15.6.1070

Hunkin, N. M., Squires, E. J., Parkin, A. J., & Tidy, J. A. (1998). Are the benefits of errorless learning dependent on implicit memory? *Neuropsychologia, 36*(1), 25–36. doi: http://dx.doi.org/10.1016/S0028-3932(97)00106-1

Kessels, R. P. C., & de Haan, E. H. F. (2003). Implicit learning in memory rehabilitation: A meta-analysis on errorless learning and vanishing cues methods. *Journal of Clinical and Experimental Neuropsychology, 25*(6), 805–814. doi: 10.1076/jcen.25.6.805.16474

Lappan, G., Fey, J. T., Fitzgerald, W. M., Friel, S. N., & Phillips, E. D. (1998). *Variables and patterns: Introducing algebra* (Student ed.). Palo Alto, CA: Dale Seymour Publications.

Larson, R., Boswell, L., Kanold, T. D., & Stiff, L. (2007). *McDougal Littell math: Course 1*. Evanston, IL: McDougal, Littell.

Leschensky, W., Malloy, C., Price, J., Rath, J., & Alban, Y. (1999). *Pre-Algebra : an integrated transition to algebra & geometry* (Teacher's wraparound ed.). Westerville, OH: Glencoe/McGraw-Hill.

Leys, C., Ley, C., Klein, O., Bernard, P., & Licata, L. (2013). Detecting outliers: Do not use standard deviation around the mean, use absolute deviation around the median. *Journal of Experimental Social Psychology, 49*(4), 764–766. doi: http://dx.doi.org/10.1016/j.jesp.2013.03.013

Litman, L., & Reber, A. S. (2005). Implicit cognition and thought. In K. J. Holyoak & R. G. Morrison (Eds.), *The Cambridge handbook of thinking and reasoning.* (pp. 431–453). New York: Cambridge University Press.

Maxwell, J. P., Masters, R. S. W., Kerr, E., & Weedon, E. (2001). The implicit benefit of learning without errors. *The Quarterly Journal of Experimental Psychology A: Human Experimental Psychology, 54A*(4), 1049–1068. doi: 10.1080/02724980143000073

Mayer, R. E., & Moreno, R. (1998). A split-attention effect in multimedia learning: Evidence for dual processing systems in working memory. *Journal of Educational Psychology, 90*(2), 312–320. doi: 10.1037/0022-0663.90.2.312

Moreno, R., & Mayer, R. E. (1999). Cognitive principles of multimedia learning: the role of modality and contiguity. *Journal of Educational Psychology, 91*(2), 358–368. doi: 10.1037/0022-0663.91.2.358

National Center for Education, S. (2011). The Nation's Report Card: Mathematics 2011. National Assessment of Educational Progress at Grades 4 and 8. NCES 2012–458: National Center for Education Statistics.

National Council of Teachers of Mathematics. (2000). *Principles and standards for school mathematics*. Reston, VA: Author.

Nissen, M. J. (1992). Procedural and declarative learning: Distinctions and interactions. In L. R. Squire & N. Butters (Eds.), *Neuropsychology of* *memory* (2nd ed., pp. 203–210). New York: Guilford Press.

Nissen, M. J., & Bullemer, P. (1987). Attentional requirements of learning: Evidence from performance measures. *Cognitive Psychology, 19*(1), 1–32. doi: 10.1016/0010-0285(87)90002-8

Nissen, M. J., Knopman, D. S., & Schacter, D. L. (1987). Neurochemical dissociation of memory systems. *Neurology, 37*(5), 789–794.

Page, M., Wilson, B. A., Shiel, A., Carter, G., & Norris, D. (2006). What is the locus of the errorless-learning advantage? *Neuropsychologia, 44*(1), 90–100. doi: http://dx.doi.org/10.1016/j.neuropsychologia.2005.04.004

Parkin, A. J. (1997). The development of procedural and declarative memory. In N. Cowan (Ed.), *The development of memory in childhood* (pp. 113–137). Hove, UK: Psychology Press/Erlbaum (UK) Taylor & Francis.

Poolton, J. M., Masters, R. S. W., & Maxwell, J. P. (2005). The relationship between initial errorless learning conditions and subsequent performance. *Human Movement Science, 24*(3), 362–378. doi: 10.1016/j.humov.2005.06.006

Reber, A. S. (1967). Implicit learning of artificial grammars. *Journal of Verbal Learning & Verbal Behavior, 6*(6), 855–863. doi: 10.1016/S0022-5371(67)80149-X

Reber, A. S., Walkenfeld, F. F., & Hernstadt, R. (1991). Implicit and explicit learning: Individual differences and IQ. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 17*(5), 888–896. doi: 10.1037/0278-7393.17.5.888

Reber, P. J., & Kotovsky, K. (1997). Implicit learning in problem solving: The role of working memory capacity. *Journal of Experimental Psychology: General, 126*(2), 178–203. doi: 10.1037/0096-3445.126.2.178

Reber, P. J., & Squire, L. R. (1998). Encapsulation of implicit and explicit memory in sequence learning. *Journal of Cognitive Neuroscience, 10*(2), 248–263. doi: 10.1162/089892998562681

Renkl, A., & Atkinson, R. K. (2007). An example order for cognitive skill acquisition. In F. E. Ritter, J. Nerb, E. Lehtinen, & T. M. O'Shea (Eds.), *In order to learn: How the sequence of topics influences learning.* (pp. 95–105). New York: Oxford University Press.

Renkl, A., Atkinson, R. K., & Grosse, C. S. (2004). How Fading Worked Solution Steps Works - A Cognitive Load Perspective. *Instructional Science, 32*(1/2), 59–82. doi: 10.1023/B:TRUC.0000021815.74806.f6

Riley, G. A., Sotiriou, D., & Jaspal, S. (2004). Which is more effective in promoting implicit and explicit memory: The method of vanishing cues or errorless learning without fading? *Neuropsychological Rehabilitation, 14*(3), 257–283. doi: 10.1080/09602010343000057

Saxon, J. H. (2003). *Algebra 1 : An incremental development*. Norman, OK: Saxon Publishers.

Schneider, W., Eschman, A., & Zuccolotto, A. (2002). E-Prime user's guide. Pittsburgh: Psychology Software Tools, Inc.

Schwartz, D. L., & Bransford, J. D. (1998). A time for telling. *Cognition and Instruction, 16*(4), 475–522. doi: 10.1207/s1532690xci1604_4

Tailby, R., & Haslam, C. (2003). An investigation of errorless learning in memory-impaired patients: Improving the technique and clarifying theory. *Neuropsychologia, 41*(9), 1230–1240. doi: http://dx.doi.org/10.1016/S0028-3932(03)00036-8

Warmington, M., Hitch, G. J., & Gathercole, S. E. (2013). Improving word learning in children using an errorless technique. *Journal of Experimental Child Psychology, 114*(3), 456–465. doi: http://dx.doi.org/10.1016/j.jecp.2012.10.007

Wessel, J. R., Haider, H., & Rose, M. (2012). The transition from implicit to explicit representations in incidental learning situations: More evidence from high-frequency EEG coupling. *Experimental Brain Research, 217*(1), 153–162. doi: 10.1007/s00221-011-2982-7

Willingham, D. B., & Goedert-Eschmann, K. (1999). The relation between implicit and explicit learning: Evidence for parallel development. *Psychological Science, 10*(6), 531–534. doi: 10.1111/1467-9280.00201

Willingham, D. B., Nissen, M. J., & Bullemer, P. (1989). On the development of procedural knowledge. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15*(6), 1047–1060. doi: 10.1037/0278-7393.15.6.1047

Woltz, D. J. (1988). An investigation of the role of working memory in procedural skill acquisition. *Journal of Experimental Psychology: General, 117*(3), 319–331. doi: 10.1037/0096-3445.117.3.319